P1 1170459

BEST AVAILABLE COPY

# THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

**UNITED STATES DEPARTMENT OF COMMERCE**
**United States Patent and Trademark Office**

**May 18, 2004**

**THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.**

**APPLICATION NUMBER:** *60/458,105*
**FILING DATE:** *March 26, 2003*
**RELATED PCT APPLICATION NUMBER:** *PCT/US04/09376*

REC'D 21 MAY 2004
WIPO PCT

**By Authority of the**
**COMMISSIONER OF PATENTS AND TRADEMARKS**

M. Sias

**M. SIAS**
**Certifying Officer**

**PRIORITY DOCUMENT**
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

03-28-03 60458105 .032603 PRO

03/26/03

JC 6 U.S. PTO

60/458105 U.S. PTO

Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

# PROVISIONAL APPLICATION COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION under 37 CFR 1.53(c).

DOCKET NUMBER: H00498.70198.US
Express Mail Label No. EV 208516365US
Date of Deposit: March 26, 2003

INVENTOR(S)/APPLICANT(S)

| LAST NAME | FIRST NAME | MIDDLE INITIAL | RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY) |
|---|---|---|---|
| Kirschner | Marc | W. | U.S.A. |
| Jebanathirajah | Judith | | U.S.A. |

[X] ***Additional inventors are being named on the separately numbered sheets attached hereto.***

TITLE OF THE INVENTION (280 characters max)

**DETECTION OF INTERACTIONS OF PROTEINS AND OTHER MOLECULES USING MASS SPECTROSCOPY**

*CORRESPONDENCE ADDRESS*

**CUSTOMER NUMBER:** **23628**

ENCLOSED APPLICATION PARTS *(check all that apply)*

[X ]Specification - Number of Pages = 55

[X ]Drawing(s) - Number of Sheets = 8

[X ]Application Data Sheet, See 37 CFR 1.76

[X ]Return receipt postcard

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.
[ ] No
[X] Yes, the name of the U.S., Government Agency and the Government Contract Number are:
This invention was sponsored by the NIH, Grant Nos. R01GM39023 and R01HD37277.

[ ] Other:

METHOD OF PAYMENT *(check all that apply)*

[ ] A check is enclosed to cover the Provisional Filing Fees.

[X] The Commissioner is hereby authorized to charge the Provisional Filing Fee and any additional fees or credit overpayment to Deposit Account 23/2825. A duplicate of this sheet is enclosed.

[ ]Small Entity Status is claimed.

PROVISIONAL FILING FEE AMOUNT $ 160.00

Respectfully submitted,

March 26, 2003
Date

Michael J. Pomianek, Reg. No. 46,190
Timothy J. Oyer, Ph.D., Reg. No. 36,628
Telephone No.: 617-720-3500

# PROVISIONAL APPLICATION COVER SHEET

## Additional Page

ATTORNEY DOCKET NUMBER: H00498.70198.US
Express Mail Label No. EV 208516365US
Date of Deposit: March 26, 2003

INVENTOR(S)/APPLICANT(S)

| LAST NAME | FIRST NAME | MIDDLE INITIAL | RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY) |
|---|---|---|---|
| Yousaf | Muhammad | N. | U.S.A. |

# DETECTION OF INTERACTIONS OF PROTEINS AND OTHER MOLECULES USING MASS SPECTROSCOPY

## FEDERALLY SPONSORED RESEARCH

This invention was sponsored by the NIH, Grant Nos. R01GM39023 and R01HD37277. The Government may have certain rights to this invention.

## BACKGROUND

### Field of Invention

This invention generally relates to proteins and small molecule arrays on self-assembled monolayers, and in particular, to the detection of proteins and small molecule arrays on self-assembled monolayers using mass spectrometry.

### Discussion of Related Art

Protein microarrays are capable of analyzing thousands of samples in short time periods. However, protein microarrays are generally unable to systematically identify proteins in a complex of whole cell lysate, due to factors such as ill-defined coupling chemistry and nonspecific adsorption. It can also be difficult in many cases to identify proper binding partners for the proteins in such cases.

Protein microarrays often are printed on standard glass slides that are activated for binding. The detection methods typically used include ELISA detection and fluorescently-labeled antibodies, which limit the detection of protein interactions to known proteins. Additionally, the use of fluorescent-tagged molecules to identify proteins may change the nature of the protein interactions being studied, as the fluorescent tag is typically an organic moiety that can interact in some fashion with the protein.

## SUMMARY OF INVENTION

This invention generally relates to the detection of proteins and small molecule arrays on self-assembled monolayers using mass spectrometry. The subject matter of this application involves, in some cases, interrelated products, alternative solutions to a particular problem, and/or a plurality of different uses of a single system or article.

*Express Mail Label No.: EV 208516365US*
*Date of Deposit: March 26, 2003*

In one aspect, the present invention includes a method. In one set of embodiments, the method includes a step of determining, using mass spectrometry, a species suspected of being bound to a substrate having an array of entities bonded thereto, at least one of which the species is suspected of being able to bind to at least one of the entities.

In another set of embodiments, the method includes providing a substrate having an array of entities bonded thereto, allowing a species to bind to at least one of the entities, and without substantially desalting the substrate, applying mass spectrometry thereto.

In another aspect, the present invention includes an article. In one set of embodiments, the article includes a species bound to at least one of an array of entities on a substrate, wherein the species bound to the entity is detectable using mass spectrometry.

In another aspect, the present invention is directed to a method of making one or more of the embodiments described herein. In yet another aspect, the present invention is directed to a method of using one or more of the embodiments described herein.

Other advantages and novel features of the invention will become apparent from the following detailed description of various non-limiting embodiments of the invention when considered in conjunction with the accompanying figures. In cases where the present specification and a document incorporated by reference include conflicting disclosure, the present specification shall control.

## BRIEF DESCRIPTION OF DRAWINGS

Non-limiting embodiments of the present invention will be described by way of example with reference to the accompanying figures, which are schematic and are not intended to be drawn to scale. In the figures, each identical or nearly identical component illustrated is typically represented by a single numeral. For the purposes of clarity, not every component is labeled in every figure, nor is every component of each embodiment of the invention shown where illustration is not necessary to allow those of ordinary skill in the art to understand the invention. In the figures:

Fig. 1 illustrates one example method of the invention;

Fig. 2 illustrates an embodiment of the invention used to bind a protein in a cell lysate;

Fig. 3 illustrates an embodiment of the invention where a single protein is determined using mass spectrometry;

Fig. 4 illustrates an embodiment of the invention where a protein fragment is determined using mass spectrometry;

Fig. 5 illustrates an embodiment of the invention where a protein-protein interaction is determined using mass spectrometry;

Fig. 6 illustrates additional mass spectrometry data, in accordance with an embodiment of the invention;

Fig. 7 illustrates one embodiment of the invention, showing a self-assembled monolayer on a microarray; and

Fig. 8 illustrates a binding partner attached to a substrate, in accordance with one embodiment of the invention.

## DETAILED DESCRIPTION

This invention in certain embodiments relates to the determination of proteins and small molecule arrays on self-assembled monolayers using mass spectrometry. In one set of embodiments, the invention relates to methods for determining proteins and small molecules bound to self-assembled monolayers using mass spectroscopy techniques such as MALDI and MALDI TOF techniques. This combination allows, for example, the systematic identification of unknown proteins from complex whole cell lysates. Identification of novel interactions can be achieved in some cases in instances where the binding partner to a particular target species is unknown. In another embodiment, the invention relates to methods of attaching a species to self-assembled monolayers on substrates such that the substrate can be used in a mass spectrometer without requiring additional exposure of the substrate to water. For example, a target species could be detected and/or analyzed using mass spectrometry in the presence of similar contaminating species, without first removing the contaminating species in some fashion.

Each of the following is incorporated herein by reference in its entirety: U.S. Patent No. 5,512,131, issued 04/30/96 to Kumar, et al.; International Patent Publication No. WO 96/29629, published 06/26/96, by Whitesides, et al.; International Patent Publication No. WO 99/54786, published 10/28/99, by Jackman, et al.; International Patent Publication No. WO 01/70389, published 09/27/01, by Ostuni, et al.

The term "determining," as used herein, generally refers to the analysis of a substance, for example, quantitatively or qualitatively, or the detection of the presence or

absence of the substance. "Determining" may also refer to the analysis of an interaction between two or more substances, for example, quantitatively or qualitatively, or by detecting the presence or absence of the interaction. In one aspect of the invention, the determination of a substance includes using mass spectrometry, as further discussed below.

The microarrays of the present invention may be formed on a substrate or support. The characteristics of these substrates may vary widely, depending upon the intended use. The microarray substrate may be formed in essentially any shape. In one set of embodiments, the substrate has at least one surface which is substantially planar. However, in some cases, the substrate may also include indentations, protuberances, steps, ridges, terraces and the like. In certain instances, the substrate can be in the form of a concave surface, a convex surface, a disc, a tubing, a cone, a sphere, and/or other geometric forms. In some cases, several substrate surfaces may be combined in some fashion. For example, two flat substrate surfaces with microarrays formed on both surfaces according to the invention may be joined together.

Suitable substrate materials include, but are not limited to, glasses, ceramics, plastics, metals, alloys, carbon, papers, agarose, silica, quartz, cellulose, polyacrylamide, polyamide, polyimide, and gelatin, as well as other polymer supports, other solid-material supports, or flexible membrane supports. Polymers that may be used in the substrate include, but are not limited to, polystyrene, poly(tetra)fluoroethylene (PTFE), polyvinylidenedifluoride, polycarbonate, polymethylmethacrylate, polyvinylethylene, polyethyleneimine, polyoxymethylene (POM), polyvinylphenol, polylactides, polymethacrylimide (PMI), polyalkenesulfone (PAS), polypropylene, polyethylene, polyhydroxyethylmethacrylate (HEMA), polydimethylsiloxane, polyacrylamide, polyimide, various block co-polymers, etc. The substrate may also comprise a combination of materials, optionally water-permeable, in some embodiments. In one embodiment, the substrate is disposable. In another embodiment, the substrate is reusable.

In some cases, the surface of the substrate may be modified in some fashion, for example, to create suitable reactive groups. Such reactive groups may include, for instance, simple chemical moieties such as amino, hydroxyl, carboxyl, carboxylate, aldehyde, ester, ether (e.g. thio-ether), amide, amine, nitrile, vinyl, sulfide, sulfonyl, phosphoryl, or similar chemically reactive groups. The reactive groups may also comprise more complex moieties that include, but are not limited to, maleimide, N-hydroxysuccinimide, sulfo-N-

hydroxysuccinimide, nitrilotriacetic acid, activated hydroxyl, haloacetyl (e.g., bromoacetyl, iodoacetyl), activated carboxyl, hydrazide, epoxy, aziridine, sulfonylchloride, trifluoromethyldiaziridine, pyridyldisulfide, N-acyl-imidazole, imidazolecarbamate, vinylsulfone, succinimidylcarbonate, arylazide, anhydride, diazoacetate, benzophenone, isothiocyanate, isocyanate, imidoester, fluorobenzene, biotin and avidin. The reactive groups may also be, for instance, silanes, Si-OH, silicon oxide, silicon nitride, primary amines or aldehyde groups. Techniques of placing such reactive groups on a substrate by mechanical, physical, electrical or chemical means are well known in the art. In some cases, the substrate may first be treated with reagents such as a strong acid before reactive groups are created on the surface. One of ordinary skill in the art will be able to readily identify suitable techniques for modifying the surface of the substrate as necessary for a particular application.

The substrate may be conductive in some embodiments. Conductive substrates may be useful, for example, to facilitate desorption of analytes from the substrate, for example, during mass spectroscopy analysis of binding partners bound to the substrate.

Aldehyde groups may be useful in some embodiments to attach proteins to the substrate, as the aldehyde groups may react with N-termini of the proteins, thus allowing the proteins to interact with other proteins and/or small molecules in solution.

In some cases, a buffer containing bovine serum albumin (BSA), casein, and/or nonfat milk may be applied to the substrate to block later non-specific binding between analytes and unreacted groups on the slide.

Linker molecules optionally may be added to the surface of the substrate to make the surface suitable for further attachment chemistry, for example, for attachment of self-assembled monolayers, as further discussed below.

As used herein, the term "linker" means a chemical moiety which can immobilize a binding partner with respect to the substrate, for example, by use of reactive groups on the substrate. As used herein, a component that is "immobilized relative to" or "bonded to" another component either is fastened to the other component or is indirectly fastened to the other component, e.g., by being fastened to a third component to which the other component also is fastened. For example, a binding partner is immobilized relative to a substrate if an entity (such as a linker) fastened to the binding partner attaches to the surface, where the entity can be a single entity, a complex entity of multiple species, etc. In certain embodiments, a component that is immobilized relative to another component is immobilized

using bonds that are stable, for example, in solution or suspension. In some embodiments, non-specific binding of a component to another component, where the components may easily separate due to solvent or thermal effects, is not preferred.

Linkers may be selected from any suitable class of compounds and may comprise polymers or copolymers of organic acids, aldehydes, alcohols, thiols, amines and the like. In one aspect, the linker may be a self-assembled monolayer, as further described below. As additional examples, the linker may include polymers or copolymers of hydroxy-, amino-, or di-carboxylic acids, such as glycolic acid, lactic acid, sebacic acid, or sarcosine. Alternatively, polymers or copolymers of saturated or unsaturated hydrocarbons such as ethylene glycol, propylene glycol, saccharides, and the like may be employed as linkers in certain embodiments. In some cases, the linker may be of an appropriate length that allows the binding partner to interact freely with molecules in a sample solution and to form effective binding.

In one set of embodiments, the linker may include a self-assembled monolayer-forming species. As used herein, the term "self-assembled monolayer" (SAM) refers to a relatively ordered assembly of molecules chemisorbed on a surface, in which the molecules are oriented approximately parallel to each other and roughly perpendicular to the surface. Each of the molecules includes a functional group that can be attached to the surface of the substrate, and a portion that interacts with neighboring molecules in the monolayer to form the relatively ordered array. Some of the methods that can be used to form a SAM are described in U.S. Patent No. 5,620,850, which is hereby incorporated by reference. See also, for example, Laibinis, P. E., Hickman, J., Wrighton, M. S., Whitesides, G. M., *Science*, 245:845 (1989); Bain, C., Evall, J., Whitesides, G. M., *J. Am. Chem. Soc.*, 111:7155-7164 (1989); Bain, C., Whitesides, G. M., *J. Am. Chem. Soc.*, 111:7164-7175 (1989), each of which is incorporated herein by reference.

Certain embodiments of the invention make use of self-assembled monolayers (SAMs) attached to surfaces of microarrays. In one set of embodiments, SAMs formed essentially completely of synthetic molecules may be used in at least a portion of a microarray. "Synthetic molecule," in this context, means a molecule that is not naturally occurring, rather, one synthesized under the direction of human or human-created or human-directed control. In some cases, the SAM can be made up of SAM-forming species that form close-packed SAMs at surfaces, and/or these species in combination with other species able

to participate in a SAM. In some embodiments, some of the species that participate in the SAM include a functionality that binds, optionally covalently, to the surface, such as a thiol which will bind to a gold surface covalently.

In one set of embodiments, the SAMs on the surface may include an $EG_3$-terminated thiol, $HS-(CH_2)_{11}-(O-CH_2-CH_2)_3-OH$ ("$C_{11}EG_3$") were used. $EG_3$-terminated SAMs on the surface are able to resist the adsorption of proteins. The $EG_3$-terminated SAMs may then surround regions that proteins will bind to, for example, as shown in Fig. 1.

In some cases, the spots on the substrate may be associated with a self-assembled monolayer ("SAM"). For example, the island(s) and/or spot(s) of the array can be defined by SAMs. In some cases, a self-assembled monolayers may be defined by molecules each having a functional group that selectively attaches to a particular surface, the remainder of most or all of the molecules having characteristics allowing them to interact with neighboring molecules in the monolayer to form a relatively ordered array. The molecules are generally largely or completely organic. Monolayers may be produced with varying characteristics and with various functional groups at the free end of the molecules which form the SAM (direction away from the surface to which the SAM attaches). Thus, SAMs may be formed which are generally hydrophobic or hydrophilic, generally cytophobic or cytophilic, or generally biophobic or biophilic. Additionally, SAMs with these or other characteristics can be formed and then modified to expose different functionalities. And SAMs with very specific binding affinities can be produced. This allows for the production of patterned SAMs which will adhere one or more protein and/or small molecules in specific and predetermined patterns.

One non-limiting example of micropatterned SAMs and their application is described in U.S. Patent No. 5,776,748, issue July 7, 1998, entitled "Method of Formation of Microstamped Patterns on Plates for Adhesion of Cells and other Biological Materials, Devices and Uses Therefor," by Singhvi, *et al.*, incorporated herein by reference.

A self-assembled monolayer on a surface of a microarray, in accordance with the invention, may be comprised of a mixture of species (e.g. thiol species when the colloid has a gold surface) that can present (expose) essentially any chemical or biological functionality. For example, such species can include tri-ethylene glycol-terminated species (e.g. tri-ethylene glycol-terminated thiols) to resist non-specific adsorption, and other species (e.g. thiols) terminating in a binding partner of an affinity tag, e.g. terminating in a chelate that can

coordinate a metal such as nitrilotriacetic acid which, when in complex with nickel atoms, captures a metal binding tagged-species such as a histidine-tagged binding species. In some embodiments of the invention, a self-assembled monolayer is formed on substrates coated with gold or silver.

In some cases, a binding partner may be attached to the linker. The binding partner may be any of a variety of different types of naturally occurring or synthetic molecules, including those having biological significance ("biomolecules"). In one set of embodiments, the binding partner is a small molecule, a peptide, or a protein. As other examples, the binding partner may include naturally occurring molecules or molecule fragments such as nucleic acids, nucleic acid analogs (e.g., peptide nucleic acid), polysaccharides, phospholipids, capture proteins including glycoproteins, peptides, enzymes, cellular receptors, immunoglobulins, antigens, naturally occurring ligands, polymers, and antibodies (including antibody fragments) such as antigen-binding fragments (Fabs), Fab' fragments, pepsin fragments ($F(ab')_2$ fragments), scFv, Fv fragments, single-domain antibodies, dsFvs, Fd fragments, and diabodies, full-length polyclonal or monoclonal antibodies, and antibody-like fragments, such as modified fibronectin, CTL-A4, and T cell receptors.

The term "binding partner" refers to a molecule that can undergo binding with a particular molecule. For example, the binding partner may a protein, a peptide, or a small molecule. "Proteins" and "peptides" are well-known terms in the art, and are not precisely defined in the art in terms of the number of amino acids that each includes. As used herein, these terms are given their ordinary meaning in the art. Generally, peptides are amino acid sequences of less than about 100 amino acids in length, but can include sequences of up to 300 amino acids. Proteins generally are considered to be molecules of at least 100 amino acids. "Small molecule", as used herein, means a molecule less than 5 kilodalton, more typically less than 1 kilodalton. As used herein, "small molecule" excludes proteins. In some cases, the binding partner may be selected from molecular libraries, such as those commercially available. A "dalton" (Da) is an alternate name for the unified atomic mass unit (grams/mole). Similarly, a "kilodalton" (kDa) is 1000 daltons.

Biological binding partners are examples of binding partners. For example, Protein A is a binding partner of the biological molecule IgG, and vice versa. The term "binding" refers to the interaction between a corresponding pair of molecules or surfaces that exhibit mutual affinity or binding capacity, typically due to specific or non-specific binding or interaction,

including, but not limited to, biochemical, physiological, and/or chemical interactions. "Biological binding" defines a type of interaction that occurs between pairs of molecules including proteins, nucleic acids, glycoproteins, carbohydrates, hormones and the like. Specific non-limiting examples include antibody/antigen, antibody/hapten, enzyme/substrate, enzyme/inhibitor, enzyme/cofactor, binding protein/substrate, carrier protein/substrate, lectin/carbohydrate, receptor/hormone, receptor/effector, complementary strands of nucleic acid, protein/nucleic acid repressor/inducer, ligand/cell surface receptor, virus/ligand, virus/cell surface receptor, etc.

Any suitable method may be used to bind or fasten a binding partner to a linker or a substrate, such suitable methods being known in the art. As used herein, ("fastened to" or "attached to") or ("adapted to be fastened to" or "adapted to be attached to") as used in the context of a species relative to another species or a species relative to a surface of an article (such as a substrate), means that the species and/or surfaces are chemically or biochemically linked to or adapted to be linked to, respectively, each other via covalent attachment, attachment via specific biological binding (e.g., biotin/streptavidin), coordinative bonding such as chelate/metal binding, or the like. For example, "fastened" or "attached" in this context also includes multiple chemical linkages, multiple chemical/biological linkages, etc. As another example, a moiety covalently linked to a thiol is adapted to be fastened to a gold-coated surface since thiols are able to bind gold covalently. Similarly, a species carrying a metal binding tag is adapted to be fastened to a surface (e.g., the surface of a colloid) that carries a molecule covalently attached to the surface (such as thiol/gold binding), which molecule also presents a chelate coordinating a metal. A species also is adapted to be fastened to or attached to a surface if a surface carries a particular nucleotide sequence, and the species includes a complementary nucleotide sequence.

The binding may be specific or non-specific, depending on the application. As used herein, "specifically bound" or "adapted to be specifically bound" means a species is chemically or biochemically linked to or adapted to be linked to, respectively, another specimen or to a surface as described above with respect to the definitions of "fastened to," "attached to," "adapted to be fastened to," and "adapted to be attached to," but excluding essentially all non-specific binding. "Covalently fastened" means fastened via essentially nothing other than one or more covalent bonds. As one example, a species that is attached to a carboxylate-presenting alkyl thiol by essentially nothing other than one or more covalent

bonds, which is, in turn, fastened to a gold-coated surface of a substrate, is covalently fastened to that substrate.

Examples of methods of coupling the binding partner to the substrate or linker include reactions that form linkage such as thioether bonds, disulfide bonds, amide bonds, carbamate bonds, urea linkages, ester bonds, carbonate bonds, ether bonds, hydrazone linkages, Schiff-base linkages, and noncovalent linkages mediated by, for example, ionic or hydrophobic interactions. The form of reaction will depend, of course, upon the available reactive groups on both the substrate/linker and binding partner.

In one aspect, the present invention provides methods for the generation of arrays of binding elements immobilized on a substrate directly or via a linker. According to the methods of the present invention, high density microarrays, with a density over 10 spots/cm$^2$, over 20 spots/cm$^2$, over 30 spots/cm$^2$, over 50 spots/cm$^2$, over 75 spots/cm$^2$, over 100 spots/cm$^2$, or over 200 spots/cm$^2$ may be formed using the techniques described herein.

The microarrays of the invention may be produced by a number of means known to those of ordinary skill in the art. Such methods include, but are not limited to, microfluidics printing, microstamping (see, e.g., U.S. Pat. No. 5,515,131 and U.S. Pat. No. 5,731,152), microcontact printing (see, e.g., PCT Publication WO 96/29629) and inkjet head printing.

In some cases, a binding partner may be tagged in some fashion, for example, with fluorescent, radioactive, chromatic and other physical or chemical labels or epitopes. However, in other cases, the binding partner may be unlabeled, and detection may occur through the use of mass spectroscopy.

A wide variety of known mass spectrometry techniques may be used in the invention. Generally, a mass spectrometer is able to create charged particles (ions) from molecules. The ions are then accelerated by manipulation of the charged particles through the mass spectrometer, while uncharged molecules and fragments are removed. For example, manipulation may occur through the manipulation of voltages to control the path of the ions, for example, to reach a detector. Analysis of those ions provides information about molecule, for example, its molecular weight and/or structural information. There are many types of mass spectrometers and sample introduction techniques which allow a wide range of analyses.

Examples of ionization methods known to those of ordinary skill in the art include, but are not limited to, electron impact, chemical ionization, electrospray, fast atom

bombardment, and matrix-assisted laser desorption ionization ("MALDI"). MALDI may be useful in situations involving larger molecules, such as peptides, proteins, or nucleotides. Examples of mass analyzers known to those of ordinary skill in the art include, but are not limited to, quadrupole, sector (magnetic and/or electrostatic), time-of-flight (TOF), and ion cyclotron resonance (ICR). In some cases, more than one mass spectroscopy technique may be used, or the mass spectroscopy technique may be combined with other techniques, for example, as in GC/MS, LC/MS, MS/MS, or MALDI/MS. Coupling two stages of mass analysis (e.g., as in MS/MS) may be useful in certain cases in identifying compounds in complex mixtures and in determining structures of unknown substances. For example, in certain such techniques, "parent" or "precursor" ions or mixture of ions in the source region or collected in an ion trap are fragmented and then the product ions resulting from the fragmentation are analyzed in a second stage of mass analysis. Additional information for structural analysis may be obtained from such techniques.

In one set of embodiments, the invention may be used to identify novel protein-protein interactions and/or protein-small molecule interactions in a high throughput fashion, for example using a tandem combination of MALDI MS and surface chemistry, as described above. According to one embodiment of the invention, using self-assembled monolayers on gold may provide an easily adjustable surface for preparing and studying biointerfacial science using MALDI MS. This surface may allow for defined coupling chemistries, which is important in ruling out positional ambiguities when a target is found to bind to a protein. Furthermore, the specific chemistry can be tailored to alleviate problems caused by non-specific binding and maintenance of the three-dimensional structure of the proteins attached.

Matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS) is a technique in which a co-precipitate of an UV-light absorbing matrix and a biomolecule is irradiated by a short laser pulse, e.g., on the order of nanoseconds. A portion of the laser energy is absorbed by the matrix, which prevents unwanted fragmentation of the biomolecule. Those of ordinary skill in the art will be able to identify suitable materials for matrices for use with MALDI. Typically, the material is able to absorb ultraviolet light. Examples of such matrices include, but are not limited to, 2,5-dihydroxybenzoic acid, 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid), a-cyano-4-hydroxycinnamic acid, certain cinnamic acid derivatives, trans-cinnamic acid, etc.

The ionized biomolecules are accelerated (typically using an electric field) towards a detector. During flight towards the detector, different molecules are separated according to their mass to charge ratio and reach the detector at different times. In this way each molecule yields a distinct signal. The method may be used for detection and characterization of biomolecules, such as proteins, peptides, oligosaccharides and oligonucleotides, with molecular masses between 400 and 350,000 Da. It can be very sensitive, and allow the detection of low ($10^{-15}$ to $10^{-18}$ mole) quantities of sample with an accuracy of 0.1 - 0.01 %, in some cases. MALDI TOF can be used, for example, to determine protein or peptide analysis and sequences, or provide information on microheterogeneity (e.g. glycosylation) and presence of by-products. MALDI TOF also can be used for the sequencing of nucleic acids in some cases. In some cases, the invention allows for the sequencing of molecules having higher molecular weights using mass spectroscopy techniques such as MALDI TOF, for example, molecules having a molecular weight of at least about 3 kDa, at least about 5 kDa, at least about 7 kDa, at least about 10 kDa, at least about 20 kDa, at least about 30 kDa, at least about 40 kDa, or at least about 50 kDa or more.

MALDI Tandem TOF MS can be employed for protein identification directly on the surface. In order to identity proteins binding to targets on the chip, an on-chip digestion method may be used to produce peptides from the bound proteins in one aspect of the invention. In some cases, on-chip proteolytic digestion may occur without denaturing the proteins, and/or in the presence of salt. For example, in one set of embodiments, mass spectroscopy analysis of a substrate having one or more binding partners attached thereon can occur without chromatographic purification (for example, as in an LC/MS technique). In some instances, the substrate may be used in mass spectrometry without substantially removing any salts present thereon. In one set of embodiments, one or more enzymes may be used for on-chip digestions, for example, but not limited to, trypsin, chymotrypsin, pepsin, etc.

If a substrate having an array of discrete spots is used, the spots may be maintained using an elastomeric membrane. As used herein, “elastomeric” defines an elastic polymer. The membrane (or similar articles) may be made of a polymeric material, and flexible polymeric materials are preferred in some embodiments of the invention. In one set of embodiments, the membrane comprises an elastomeric material. A variety of elastomeric materials may be suitable, especially polymers of the general classes of silicone polymers,

epoxy polymers, and acrylate polymers. Epoxy polymers are characterized by the presence of a three-member cyclic ether group commonly referred to as an epoxy group, 1,2-epoxide, or oxirane. For example, diglycidyl ethers of bisphenol A may be used, in addition to compounds based on aromatic amine, triazine, and cycloaliphatic backbones. Another example includes the well-known Novolac polymers. Examples of silicone elastomers suitable for use as a membrane include those formed from precursors including the chlorosilanes such as methylchlorosilanes, ethylchlorosilanes, phenylchlorosilanes, and the like. One useful silicone elastomer is polydimethylsiloxane. The membrane may be reusable in some cases.

Applications of the invention include, but are not limited to, determining protein-protein interactions and/or protein-small molecule interactions, . In some cases, the proteins and/or small molecules to be determined may be in a mixture comprising multiple proteins and/or small molecules, for example, as would be present in a cell lysate. For example, if the microarray comprises antibodies, the microarray may be exposed to a cell lysate and various protein or other small molecules may be determined by exposing the microarray to mass spectroscopy, as previously described. In some cases, the invention allows for the determination, using mass spectrometry, of a species suspected of being bound to a substrate in the presence of similar species which would interfere with determining the species in more conventional mass spectrometry techniques.

The following examples are intended to illustrate certain aspects of certain embodiments of the present invention, but do not exemplify the full scope of the invention.

Example 1

Self-assembled monolayers were prepared on gold coated glass slides and these were used to make protein and small molecule arrays in this example. This example combines self-assembled monolayers on gold surfaces and MALDI tandem MS capabilities. This combination allowed the identification of novel protein–protein and/or protein–small molecule interactions..

The arrays were then incubated with cell lysates in order to capture binding partners. An in situ digestion using various enzymes was performed on the chip surface. These surfaces were used as MALDI targets and interrogated using mass spectrometry. The mass spectrometers used were a MALDI TOF, Voyager DE STR (Applied Biosystems,

Framingham, MA) and a MALDI TOF/TOF, ABI 4700 (Applied Biosystems). Mascot (Matrix Science, London, UK), Protein Prospector and PepSea (Protana, Odense, DK) were used to interrogate protein databases.

The spatially addressable protein chip was placed directly into the mass spectrometer such that the MALDI TOF mass spectrometer was used as a detector for proteins bound to small molecules or proteins. Intact proteins were easily detected using these chips as MALDI targets. This proved to be useful for screening combinatorial small molecule libraries for ligands using a particular protein.

In order to use this system as discovery tool, i.e., to identify proteins from complex samples such as cell lysates, binding to e.g. antibodies, an efficient on-target digestion method without subsequent washing or chromatography steps was employed. The generated peptides were sequenced/fragmented using a MALDI TOF/TOF tandem mass spectrometer providing sufficient sequence information for unambiguous protein identifications.

In one instance, an anti-CDC27 antibody was immobilized on a chip surface in order to purify the anaphase-promoting complex (APC). The chip was then incubated with HeLa cell lysate and subsequently washed. A digestion was carried out on the surface and bound proteins were identified using MALDI MS/MS. Some known members of the complex and novel binding partners were identified.

Example 2

This example describes the use of MALDI Tandem TOF mass spectrometry for the on-chip identification of protein-protein interactions in the microarray format. This example also illustrates the use of the developed methods for the detection of proteins that interact with small molecules. Further details are given in Appendix A.

The identification of the binding proteins was made possible by developing an in situ microarray proteolytic digest and by recent technological advances in MALDI tandem MS. The in situ digest produced peptides from binding proteins that were sequenced with high sensitivities, allowing the unambiguous identification of protein binding partners. Surface technology was used to reduce nonspecific surface interactions, which was advantageous as mass spectrometry is highly sensitive and any nonspecific interactions will tend to produce a high background. SAMs engineered to reduce nonspecific interactions by presenting oligo(ethyleneglycol) groups interspersed with those presenting the molecule of choice at low

densities were used on a gold surface. The use of the gold surface was advantageous for MALDI tandem MS in this example, as the ionization and acceleration of analyte molecules required that a voltage be applied to the target. The developed technology was used to identify novel and known binding partners of the Anaphase Promoting Complex (APC).

To identify proteins in this example, an on-array digest compatible with MALDI ionization was developed. An on-target digestion facilitated two methods of protein identification; by peptide mass mapping or by sequencing a peptide of interest. Both methods produced unambiguous protein identification, based on the number of proteins on the surface. Nonspecific interactions were also minimized to allow on-array target identification, as mass spectrometers are extremely sensitive and nonspecific interaction would tend to introduce a high background, thus suppressing peptide from real interactions.

In this example, a chemically flexible micro array platform was designed which allowed the immobilization of small molecules and/or proteins and which could be used as a discovery tool to identify proteins from a complex mixture binding to a particular ligand and/or novel protein interactions. In order to achieve this, several aspects of design were accounted for: reduction of nonspecific binding; developing a MALDI compatible in situ tryptic digest; engineering solutions to place the array into the mass spectrometer without compromising mass accuracy and resolution. MALDI TOF/TOF mass spectrometry was used in order to get unambiguous identification of the proteins based on peptide sequence information.

Using a MALDI TOF as a detector for the readout of the SAM-based arrays had several advantages. The gold surface – as any conductive surface – was able to quench fluorescence, thereby limiting the sensitivity of detection by fluorescence. This permitted the analysis of binding studies on small molecule and/or protein arrays due to its sensitivity and its ability to function as a simple reader/detector as well as peptide sequencing tool, useful, for example, as a discovery tool for the analysis of unknown and novel interactions.

MALDI mass spectrometry is extremely sensitive. Thus, it can be important to provide an inert surface to reduce and preferably eliminate false interactions that might take place. Certain gold surfaces were used that were selected to give specific interactions with biomolecules. For example, see U.S. Patent No. 5,776,748, incorporated herein by reference. The surfaces rely on self-assembled monolayers (SAM) that present oligo(ethyleneglycol) groups.

An example of such a technique is shown schematically in Fig. 1. Substrates were prepared by sputtering gold (12nm) on the surface of glass coverslips which were treated with titanium. Microcontact printing was used to pattern hexadecanethiol [$HS(CH_2)_{15}CH_3$] onto gold-coated surfaces. The patterns used in this example were round dots. The bare gold surfaces between the spots of alkylation were subsequently derivatized with a polyethylene glycol (PEG)-terminated alkanethiol [$HS(CH_2)_{11}$-$(OCH_2CH_2)_3OH$]. SAMs made from this alkanethiol provided inert surfaces that minimize nonspecific adsorption. As the spots were patterned in a registered and addressable manner, proteins of interest were then arrayed onto these patterns using a robotic spotter. The exposure of the hexadecanethiol substrate to a protein solution resulted in the formation of protein coated islands with a geometry and a distribution defined by the original SAM patterning. The optimal diameter of the arrayed spots for interrogation by MALDI TOF MS was experimentally determined to be 400-800 microns, separated by inert surfaces of approximately 1-2 mm. This distribution proved to be superior as it minimized spot to spot bleeding during the protein digestion and subsequent sample preparation steps.

Poly (dimethylsiloxane) (PDMS) was purchased from Dow Corning (Sylgard 184). PDMS stamps were prepared from photolithographically produced masters as described previously described in U.S. Patent No. 5,512,131, incorporated herein by reference. Hexadecanethiol was purchased from Aldrich and purified by silica gel chromatography.

Instead of utilizing pure chemisorption, as shown here, it is also feasible to array small molecules using various chemistries for covalent immobilization such as the Michael addition or the Diels-Alder reaction.

Once the small molecule or protein array was prepared, several subsequent experiments, each of which required different detectors/analytical methods for the read out, were performed. The chip was incubated a) with a single protein, b) with a mixture of known proteins, or c) with a complex uncharacterized mixture such as whole cell lysate. After incubation with the protein solution of interest the array was washed with the required stringency to remove buffer components and unbound proteins. For applications a) and b) the detection of the intact protein was performed (see below, discussion of Figure 2), whereas for application c), the generation of proteolytic peptides by on-target digestion was performed to improve the sensitivity and to facilitate peptide sequencing. The former was accomplished by adding a MALDI matrix suitable for whole protein analysis, such as dihydroxybenzoic acid

(DHB) or sinapinic acid (SA) onto the surface, and MALDI spectra was collected for the different spots on the array.

However, in cases where peptides were utilized for the analysis in order to provide unambiguous protein identification, the problem of on-target digestion without any sample carry-over from one spot to the other was solved as follows. The digest was performed in situ on the array in a humidity chamber without any prior reduction and alkylation. Subsequently, a MALDI matrix suitable for peptide analysis such as α-cyano-4-hydroxycinnamic acid was added to the digest array and the solution was evaporated under reduced pressure. The gold substrate with the array was then placed on a target that was specifically machined so as to accommodate the chip while maintaining the voltage on the surface of the chip. A gold coated array was useful for this type of analysis as the metal layer was conductive, thereby facilitating sensitive MALDI without unduly compromising accuracy and resolution.

In order to demonstrate the above described strategy with a small molecule-protein interaction, a surface was prepared with a SAM presenting a PEO-Biotin (polyethylene oxide) terminated-thiol (Figure 2c). The PEO-Biotin used had an ethylene oxide linker which renders the molecule hydrophilic and thus reduced nonspecific adsorption by proteins other than streptavidin. The chip was rinsed, dried and incubated with a 2 pmol/ml solution of FITC coupled streptavidin. This was incubated overnight on a shaker and subsequently rinsed. The fluorescence image of the chip is shown in Figure 2a. The image shows a very clean distribution of the streptavidin protein on the PEO-Biotin treated surface areas. In a following experiment, a PEO-Biotin chip incubated with streptavidin was overlaid on the chip with DHB as matrix at a concentration of 5 ug/ul. The obtained MALDI TOF mass spectrum is shown in Figure 2b. The spectrum is clean of any other nonspecifically adhering proteins present in the whole cell lysate. This demonstrates the feasibility of using this strategy to discover novel protein small molecule interactions. As this is a known protein of known mass, it was unnecessary in this example to perform a digest for more detailed analysis of the peptides.

In order to show the different approaches described above, cytochrome C was applied to a SAM-based gold chip. The protein array produced as described above was incubated with a single protein in solution. The array was washed and MALDI matrix was applied to the array on the protein coated spot as well as on the gold surface derivatized with polyethylene glycol terminated alkanes, i.e. the surface areas inert to nonspecific protein

binding. Spectra were collected by MALDI TOF MS on the spots as well as on the "inert" surface areas. The results are shown in Figure 3A. The inert surfaces were found to be completely clean of any protein whereas the spots with the physisorbed protein showed distinct signals for the singly, doubly and triply charged cytochrome C at 11564 kDa, 5782 kDa and 3855 kDa. This type of analysis, i.e. using the MALDI TOF mass spectrometer as a simple "chip reader" in order to detect the protein of interest based on its mass can be utilized in cases where a single protein was being screened against a protein, peptide or small molecule library.

However, if a complex protein mixture is used, the mass of the binding protein may not be sufficient for unambiguous identification. In these cases, an in situ digest on the target was performed in order to identify the protein by its peptide mass map or based on MS/MS data from one or several peptides. Several conditions for the in situ digestion were evaluated. The following digestion conditions were found to be advantageous: low concentrations of ammonium bicarbonate (20 mM) as volatile buffer and increased temperatures of about 50 °C (REF). MALDI MS compatible detergents at low concentrations, such as 10 mM n-octyl-glucopyranoside (OGP) or acid labile detergents (RapiGest, Waters) were used for proteins in cases where the proteins were difficult to digest. After the digestion, 0.1% trifluoroacetic acid TFA was used as wash solution to remove ammonium bicarbonate buffer and excess detergent used to help the digest. The peptides that resulted from the digest appeared to physisorb to the surface of the hydrophobic spot as did the protein and thus allowing easy washing of the chip surfaces. The high affinity of the peptides to the hydrophobic areas also improved the sensitivity of the technique by concentrating the peptides to a small sample spot diameter in the sub-millimeter size range, which in turn reduced shot-to-shot heterogeneity, as often observed for traditional MALDI sample preparations. Figure 3C shows the peptide mass fingerprint obtained after in situ digestion of cytochrome C. The digest produced numerous peptides with a sequence coverage of about 58%. After calibration using signals from trypsin autolysis peptides the spectrum was submitted for a MASCOT search, which unambiguously identified the protein as cytochrome C.

The feasibility of performing product ion experiment for sequencing purposes on a peptide on this surface at the peptide concentrations found on chip was also investigated. Using MALDI MS/MS technology including MALDI TOF/TOF and MALDI quadrupole/TOF technology permitted the use of MALDI for the detection and identification

of novel interactions. Several peptides were selected for fragmentation and a spectrum from the precursor peptide at 1169.3 Da is shown in Figure 3D. A clear sequence stretch covering most of the spectra collected were of sufficient quality to derive almost the entire peptide sequences, thus obtaining positive identifications using commonly used search algorithms such as MASCOT, Protein Prospecter and PepSea.

As it was possible to perform in situ digests and to identify a protein bound to small molecules or physisorbed to these surfaces by peptide mass fingerprinting and also by fragmenting, i.e. sequencing a peptide, the next experiments established that the SAM-based arrays could also be used for the study of protein-protein interactions including mass spectrometric identification of the binding partners. An anti-N-WASP antibody was physisorbed to a surface patterned with 400 um spots, 2 mm apart by incubating the array with a solution of the anti-N-WASP antibody. The surface was washed in order remove excess antibody prior to incubation with 1 ml of HeLa whole cell lysate spiked with 200 ug of recombinant N-WASP. The chip was subsequently extensively washed prior to overnight tryptic digestion on-target. Matrix was then added to the surface and the surface was allowed to dry under reduced pressure. The MALDI TOF mass spectrum of this digest is shown in Figure 4a. Apart from intense ion signals deriving from trypsin autolysis peptides and peptides deriving the antibody used for the experiment several other peptide ion signals are apparent which were used for a peptide mass fingerprint search after internal calibration utilizing the autolysis peaks. The MASCOT search software was used for the protein identification. The N-WASP was identified as top-scoring protein hit. As the MASCOT probability score did not provide an unambiguous identification several peptides were subjected to sequencing by MALDI tandem mass spectrometry in order to give a clear identification. The peptide at m/z 1990.2 was one of the peptides selected for sequencing by MS/MS. The acquired product ion spectrum is shown in Figure 4b. This product ion spectrum was submitted to the MASCOT search engine and retrieved as top-scoring peptide a N-WASP-derived peptide with the sequence GGPPPPPPPHSSGPPPPPAR thereby unambiguously identifying the N-WASP protein. Again as with several product ion spectra deriving from the cytochrome c digest described above, this spectrum was of good quality for manual interpretation of up peptide sequence stretches of up to 6 amino acid residues. To examine if a protein complex could be pulled down on the surface of the array an antibody to CDC27 was immobilized on the surface of the array. CDC27 is a member of the well

characterized anaphase promoting complex (APC). The human complex has 12 characterized members.

While several embodiments of the invention have been described and illustrated herein, those of ordinary skill in the art will readily envision a variety of other means and structures for performing the functions and/or obtaining the results or advantages described herein, and each of such variations or modifications is deemed to be within the scope of the present invention. More generally, those skilled in the art would readily appreciate that all parameters, dimensions, materials, and configurations described herein are meant to be exemplary and that actual par, dimensions, materials, and configurations will depend upon specific applications for which the teachings of the present invention are used. Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. It is, therefore, to be understood that the foregoing embodiments are presented by way of example only and that, within the scope of the appended claims and equivalents thereto, the invention may be practiced otherwise than as specifically described. The present invention is directed to each individual feature, system, material and/or method described herein. In addition, any combination of two or more such features, systems, materials and/or methods, if such features, systems, materials and/or methods are not mutually inconsistent, is included within the scope of the present invention.

In the claims (as well as in the specification above), all transitional phrases such as "comprising," "including," "carrying," "having," "containing," "involving," and the like are to be understood to be open-ended, i.e. to mean including but not limited to. Only the transitional phrases "consisting of" and "consisting essentially of" shall be closed or semi-closed transitional phrases, respectively, as set forth in the United States Patent Office Manual of Patent Examining Procedures, Section 2111.03.

What is claimed is:

**CLAIMS**

1. A method, comprising:

determining, using mass spectrometry, a species suspected of being bound to a substrate having an array of entities bonded thereto, at least one of which species is suspected of being able to bind to at least one of the entities.

2. The method of claim 1, wherein the determining step comprises determining the species using MALDI mass spectrometry.

3. The method of claim 2, wherein the determining step comprises determining the species using MALDI TOF mass spectrometry.

4. The method of claim 1, wherein the determining step comprises determining the species using a first mass spectrometry technique and a second mass spectrometry technique.

5. The method of claim 1, wherein the species is a small molecule.

6. The method of claim 5, wherein the small molecule is selected from a small molecule library.

7. The method of claim 1, wherein the species is a protein.

8. The method of claim 7, wherein the determining step comprises determining at least a portion of a primary sequence of the protein.

9. The method of claim 1, wherein the species is part of a mixture of substances.

10. The method of claim 9, wherein the mixture comprises cell lysate.

11. The method of claim 1, wherein the substrate includes a microarray.

12. The method of claim 1, wherein the substrate includes a protein array.

13. The method of claim 1, wherein the substrate includes a small molecule array.

14. The method of claim 1, wherein the substrate includes a self-assembled monolayer.

15. The method of claim 1, wherein the substrate includes a gold surface.

16. The method of claim 1, wherein the entities arise from a cell lysate.

17. The method of claim 1, wherein the entities are attached to the substrate .

18. The method of claim 1, wherein the entities include a protein.

19. The method of claim 1, wherein the entities include a small molecule.

20. The method of claim 1, wherein the species is able to specifically bind to at least one of the entities.

21. The method of claim 1, wherein the entities include an antibody.

22. The method of claim 1, wherein the determining step comprises determining the presence of the species.

23. The method of claim 1, wherein each entity is individually addressable by mass spectrometry.

24. A method, comprising:

providing a substrate having an array of entities bonded thereto;

allowing a species to bind to at least one of the entities; and
without substantially desalting the substrate, applying mass spectrometry thereto.

25. The method of claim 24, wherein the species includes a protein.

26. The method of claim 24, wherein the substrate includes a microarray.

27. The method of claim 24, wherein the substrate includes a protein array.

28. The method of claim 24, wherein the substrate includes a small molecule array.

29. An article, comprising:
a species bound to at least one of an array of entities on a substrate, wherein the species bound to the entity is detectable using mass spectrometry.

30. The article of claim 29, wherein the species includes a protein.

31. The article of claim 29, wherein the substrate includes a microarray.

32. The article of claim 29, wherein the substrate includes a protein array.

33. The article of claim 29, wherein the substrate includes a small molecule array.

## ABSTRACT

This invention generally relates to the determination of proteins and small molecule arrays on self-assembled monolayers using mass spectrometry. In one set of embodiments, the invention relates to methods for determining proteins and small molecules bound to self-assembled monolayers using mass spectroscopy techniques such as MALDI and MALDI TOF techniques. This combination allows, for example, the systematic identification of unknown proteins from complex whole cell lysates. Identification of novel interactions can be achieved in some cases in instances where the binding partner to a particular target species is unknown. In another embodiment, the invention relates to methods of attaching a species to self-assembled monolayers on substrates such that the substrate can be used in a mass spectrometer without requiring additional exposure of the substrate to water. For example, a target species could be detected and/or analyzed using mass spectrometry in the presence of similar contaminating species, without first removing the contaminating species in some fashion.

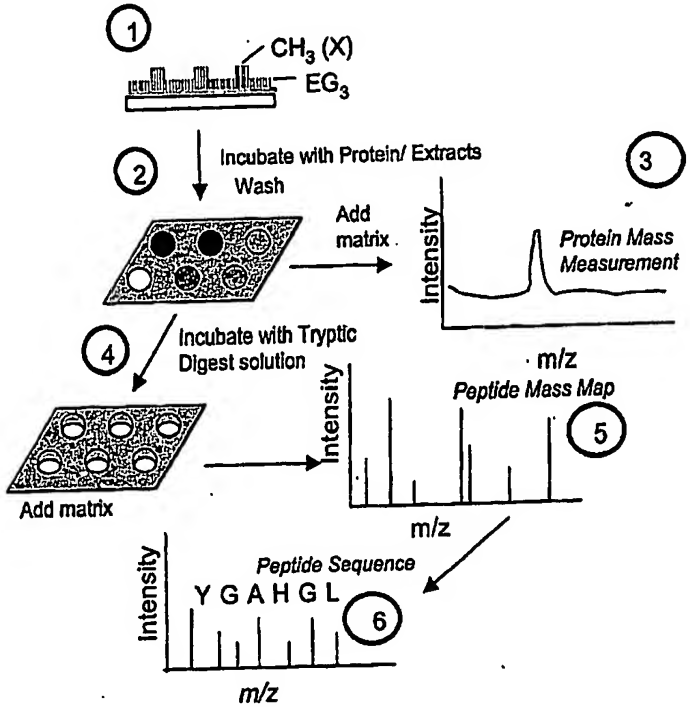



**Figure 1.**

**1.** Small molecules or hydrophobic spots are printed on the gold surface in a preregistered addressable format using microcontact printing.

**2.** The printed chip is then incubated with a protein sample of interest, this could be a single protein or a complex mixture.

**3.** The printed array is washed. A MALDI matrix suitable for protein detection can be added to the surface and the spectra can be collected at this point.

**4.** An enzymatic digest is performed on the surface by adding a solution of enzyme.

**5.** After the surface is dried, a peptide compatible matrix is then added using the same membrane. After the digest the protein can be identified using MALDI TOF mass spectrometry.

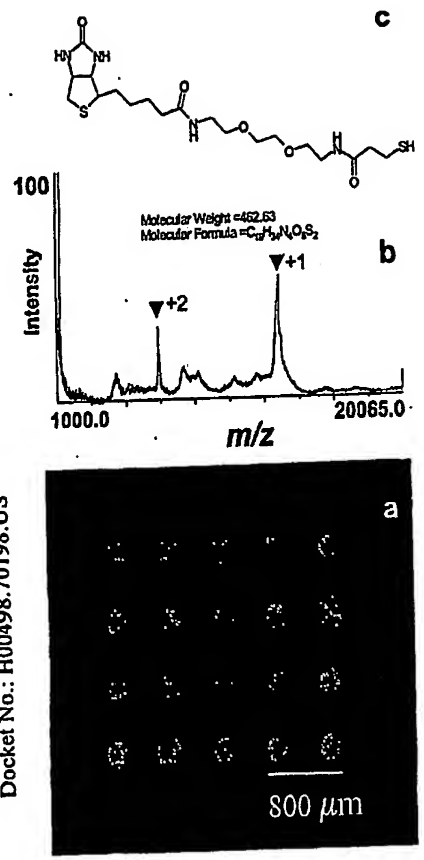



**Figure 2.** To show the feasibility of using this method with a small molecule array a biotin thiolate ligand, shown in **c** was coupled to the surface in an array format. The array was then incubated with Hela cell lysate spiked with femtomoles of FTIC coupled streptavidin. The array was then washed and imaged using a fluorescence microscope, the image is shown in panel **a** sinnapinic acid was then added to the spots on the surface and the spectrum shown in **b** was collected.

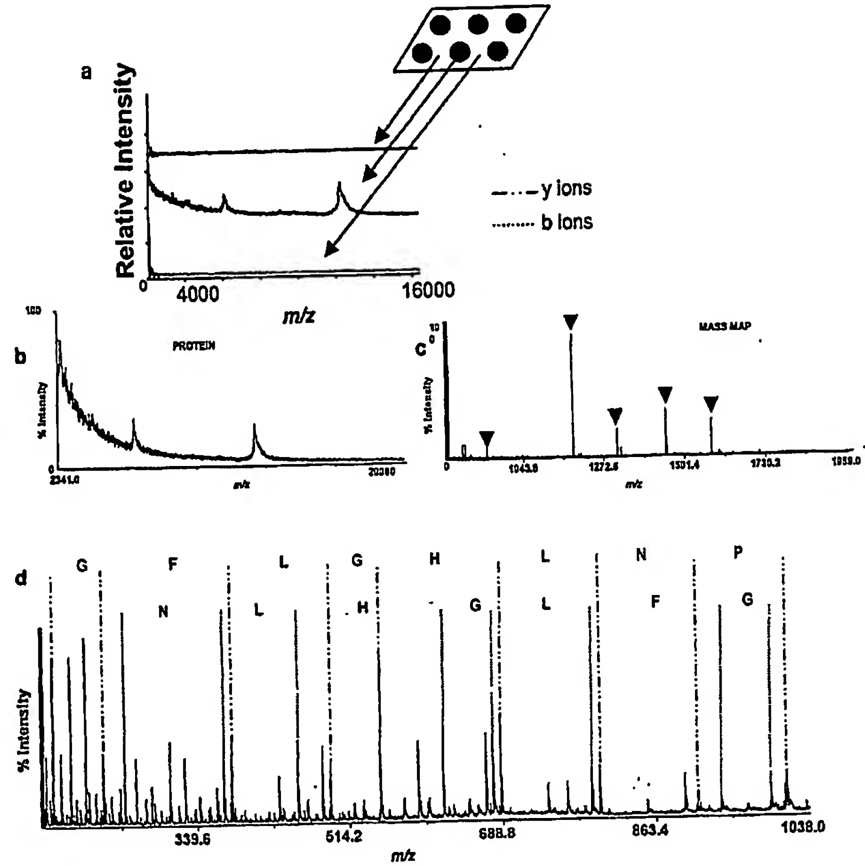



**Figure 3.a)** An incubation of the surfaces with a single protein, cytochrome C is shown here. The hydophobic spots on the chip are shown to bind the protein whereas the hydrophilic areas i.e. $EG_3$ terminated SAM show no binding of the protein. **b)** The size of the protein can be determined by mass spectrometry **c)** The peptide mass map of the protein can be obtained by performing an on chip digestion. Here we show sequence coverage of 58% **d)** If the protein mass map does not provide a postive identification the protein can be identified by sequencing a peptide giving unambiguous identification.

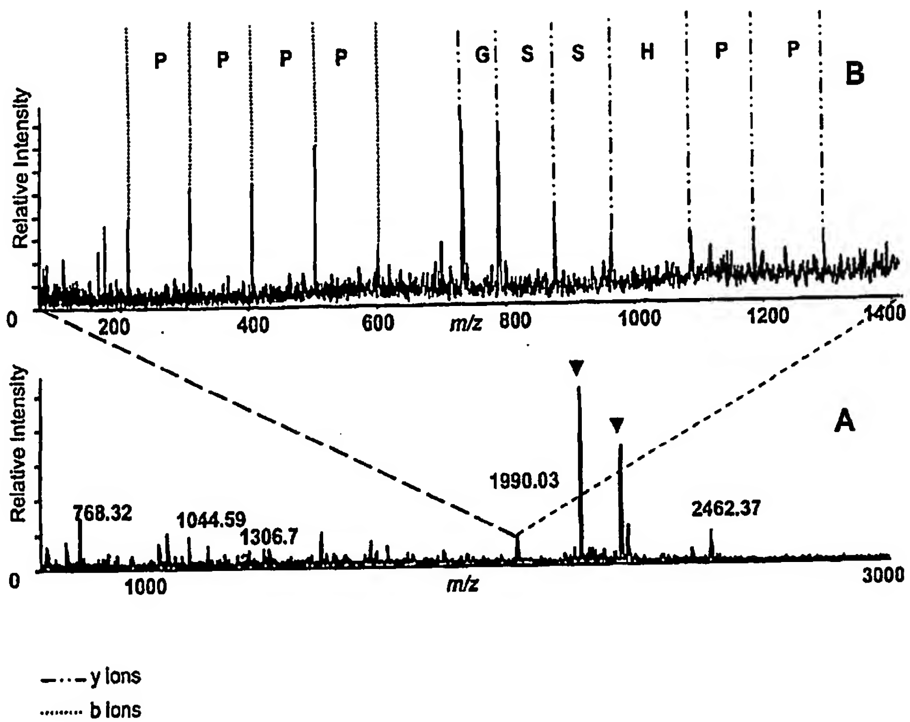



**Figure 4.** In order to investigate the feasibility of this technology for studying protein-protein interactions an antibody for N-WASP was arrayed on the chip. The chip was then incubated with whole cell lysate spiked with recombinant N-WASPand then washed. An on-chip digestion using trypsin was performed and matrix was added to the array. A peptide mass map of this chip is shown in panel A. The peptide at 1990.03 was then fragmented and identified as being GGPPPPPPPPHSSGPPPPPAR. The arrow in panel A show a signature peptide signal arising from Roche Bovine Trypsin at 2162.049 and the adjacent marked peak is 2172.1522, which on sequencing was identified as an IgG peptide.

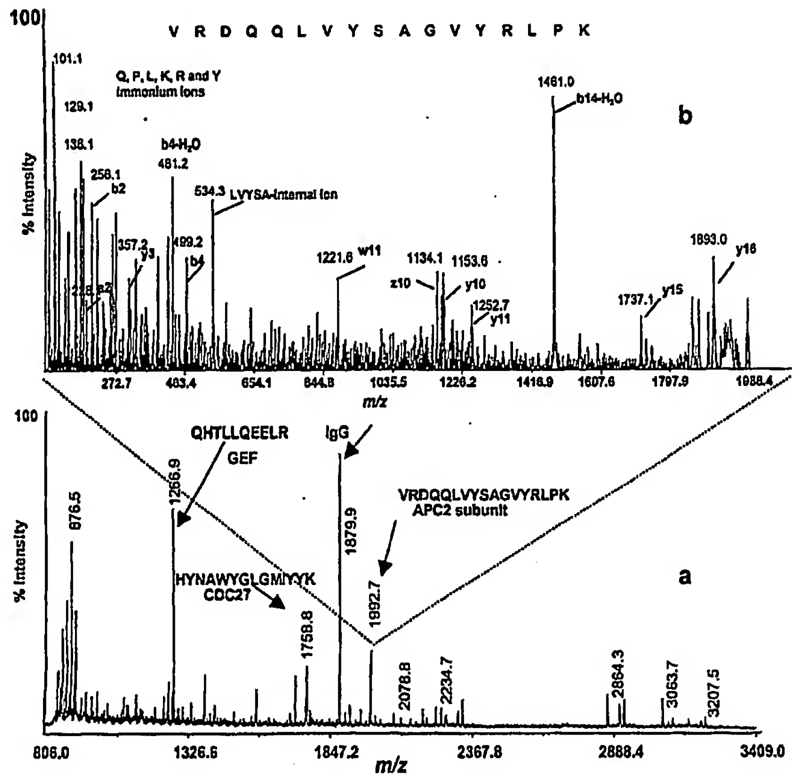



**Figure 5.** To examine the feasibility of identifying protein-protein interactions in a complex with this array format. An array with an antibody to CDC27 displayed on the surface was made. CDC27 is a member of the well characterized anaphase promoting complex (APC) also known as APC subunit 3.The array was then incubated with HeLa whole cell lysate and then washed. An in situ digest was then performed on the surface. $\alpha$-Cyano-4-hydroxycinnamic Acid was added to the surface and spectra were collected. The MALDI MS1 spectra shown in panel **a**. Some of the peptides on the surface were fragmented and data was interogated using Protein Prospector, MS-Tag, Mascot and Pepsea. Panel **b** shows a spectra which identified a peptide from APC subunit 2, an integral protein with Cullin homology.

DETECTION OF INTERACTIONS OF PROTEINS AND OTHER MOLECUES USING MASS SPECTROSCOPY
Marc W. Kirschner et al.
Serial No.: Not Yet Assigned
Docket No.: H00498.70198.US

60458105 .032603

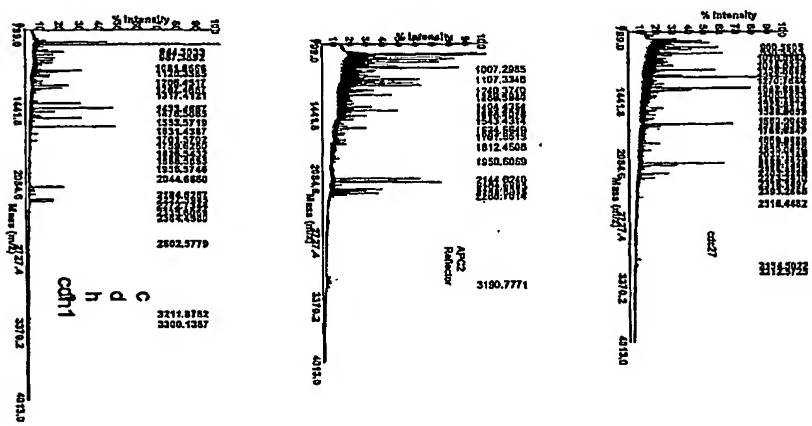


FIG. 6

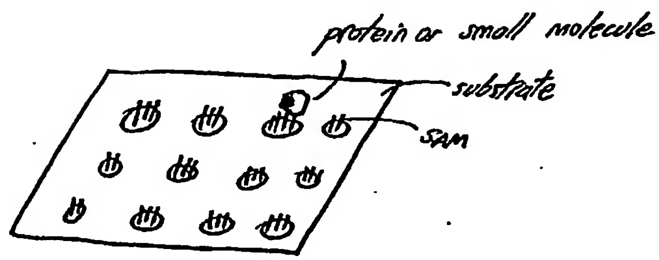



FIG. 7

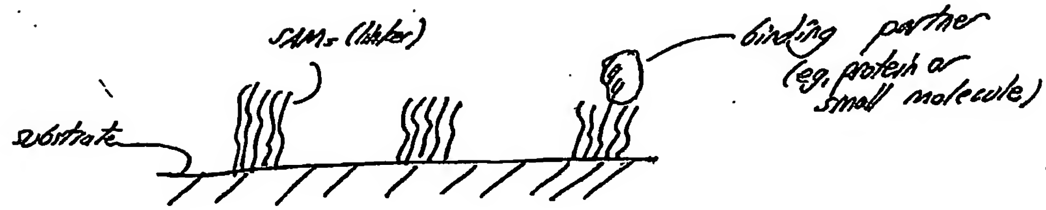

SAMs (linker)
binding partner
(eg. protein or
small molecule)
substrate

FIG. 8

Appendix A

# Progress of Modern Biochemistry

1. Identification of Activities and Functions of Proteins
2. Purification of Proteins
3. Sequence of proteins and cloning of genes
4. Studies of Protein-Protein Interactions
5. Studies of Small Molecule - Protein Interactions
6. Quantification of 4. And 5.

PRIOR ART

- **Biacore**: Measures binding energy, affinity and interactions of sample. Not high thru'put yet. Detection of interactions by Surface Plasmon Resonance.

- **Ciphergen**: Protein profiling using Reverse Phase material and mass spectrometry. No systematic protein identification.

- **Zyomyx**: Antibody capture chips, SAMs and 3-D microfluidic structures. Use Fluorescence Detection

- **BioInsights**: custom microfluidic engineering

- **Protagen, Jerini**: immobilized small molecules, peptides and proteins. Several detection methods including $^{32}P$ autoradiography.

# Existing Methods of Studying Protein-Protein Interactions

- **Yeast two hybrid system, amenable to a high thru'put approach Vidal *et al*. Proc Natl Acad Sci U S A. 1996 Sep 17;93(19):10321**

  **Limitations**: large percentage of false positives 70%, only binary interactions between protein domains studied, all interactions forced to occur in nucleus and PTM dependent interactions excluded.

- **TAP tagging followed by Gel-LCMS**, high thru'put MDS Proteomics and Cellzome. Ho et al. Nature 2002, 415:180-183, Gavin et al. Nature 2002, 415:141-47

  **Limitations:** Large amounts of sample required, labour intensive and not easily extrapolated to mammalian systems.

- **Bioinformatics**: Prediction of Protein interactions domains. Hogue, C. BMC Bioinformatics 2003 Jan 13;4(1):2

  **Limitations:** Predictions are based on known interactions and structures

# Mass Spectrometry and Protein Arrays

- What does the mass spectrometry offer to chip technology?
- The ability to detect and identify unknown partners to known targets.
- The ability to detect and identify unknown partners to unknown targets.

# Our Approach

Most Flexible
Surface Chemistry
Technology

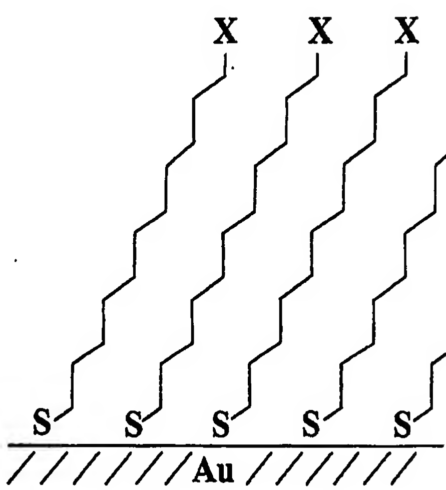
Most Versatile and
Sensitive Protein Identification
Technology

An Integrated Approach Using Self-Assembled Monolayers,
On-Chip Proteolytic Digestion and MALDI Mass Spectrometry to
Characterize Small Molecule-Protein and Protein-Protein Microarrays

# The Technology

1. Could Identify Unknown Proteins Interacting with Known or Unknown Targets
2. Would Requre Very Small Amounts of Targets and Samples (Ideally would work on the level of tissue culture experments - 250 mL of cultured cells)
3. No Background
4. High-Throughput - 1000's of samples per day
5. Synthetically Flexible for use in small molecule protein interactions
6. Low cost and widely applicable

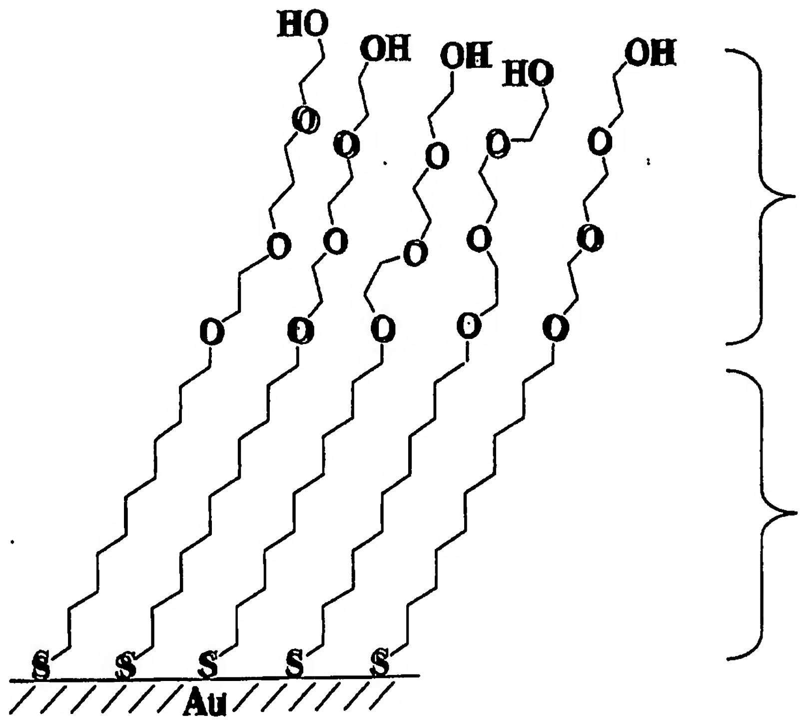
# Surface Chemistry:
## Self-Assembled Monolayers (SAMs) of Alkanethiolates on Gold




1. Can Make Mixed Monolayers
   -generate complex surfaces
2. Synthetically Flexible
   -routine organic synthesis
3. Electroactive Monolayers
   -apply potential to gold

$$RSH + Au(0)_n \longrightarrow RS^{-}Au(I) \bullet Au(0)_n + 1/2\ H_2$$

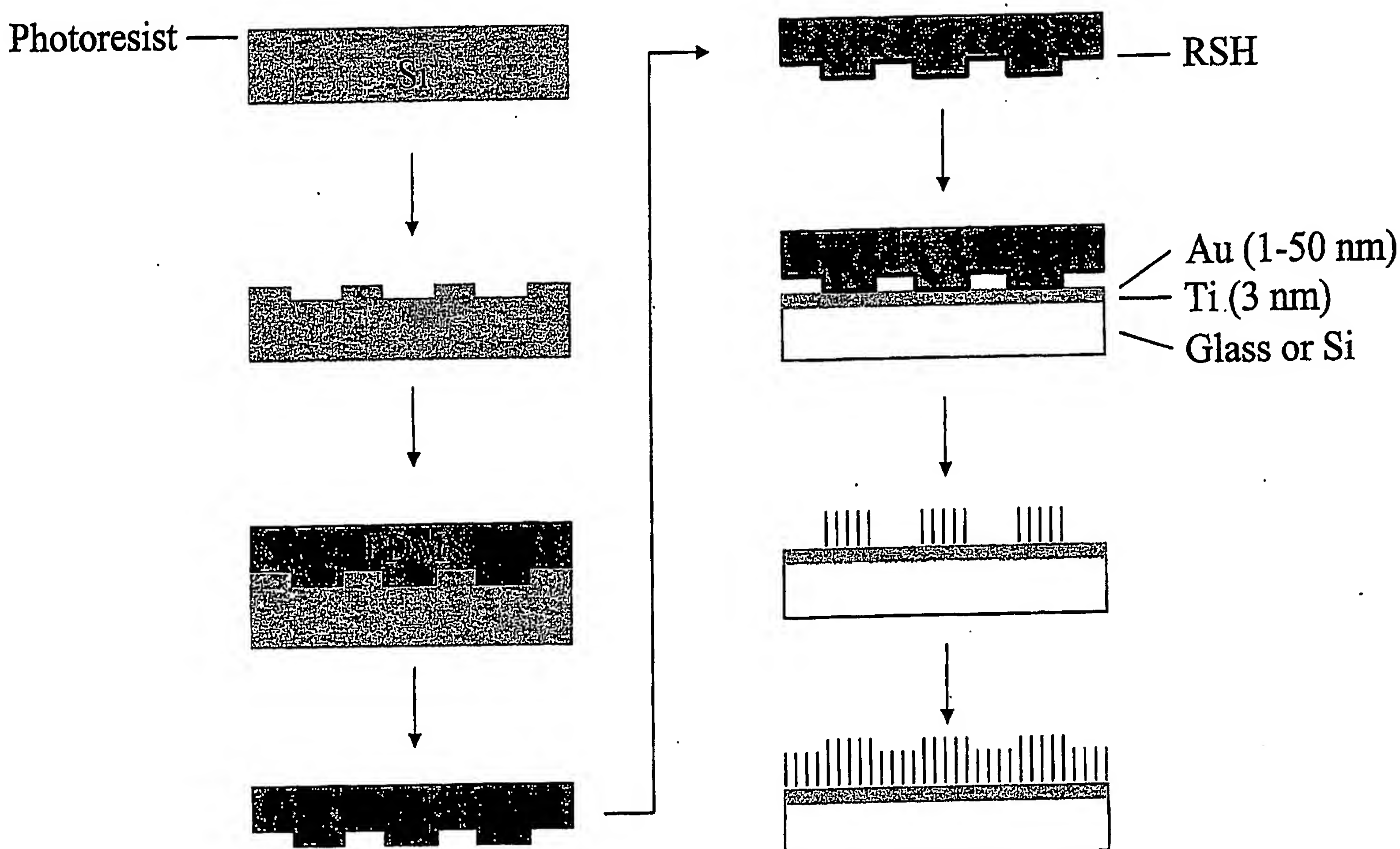
# Monolayers that Resist Adsorption of Protein




Oligo(ethylene glycol) chain
prevents protein adsorption

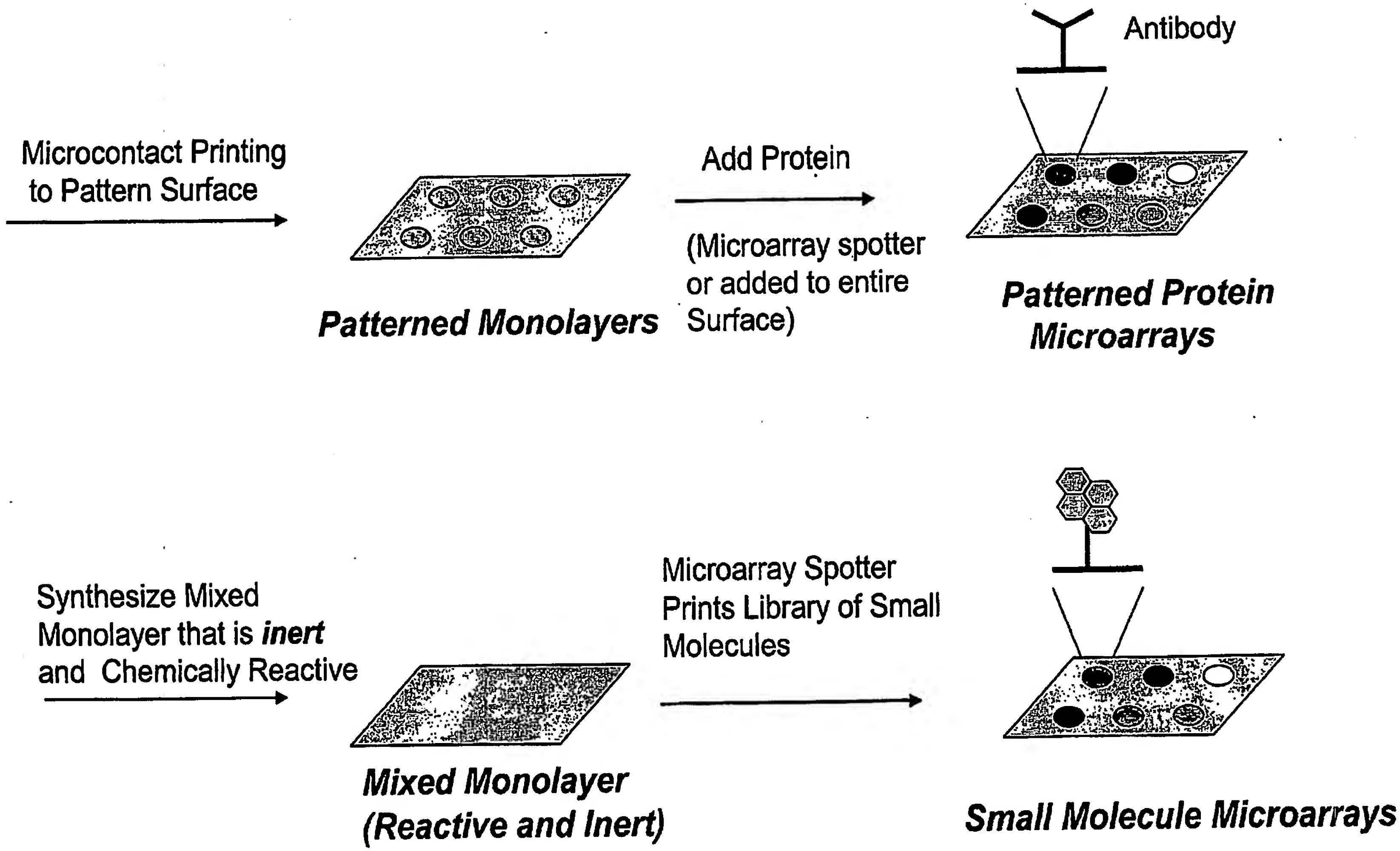
Poly(methylene) chain
gives structural order

# Patterning Monolayers with Microcontact Printing

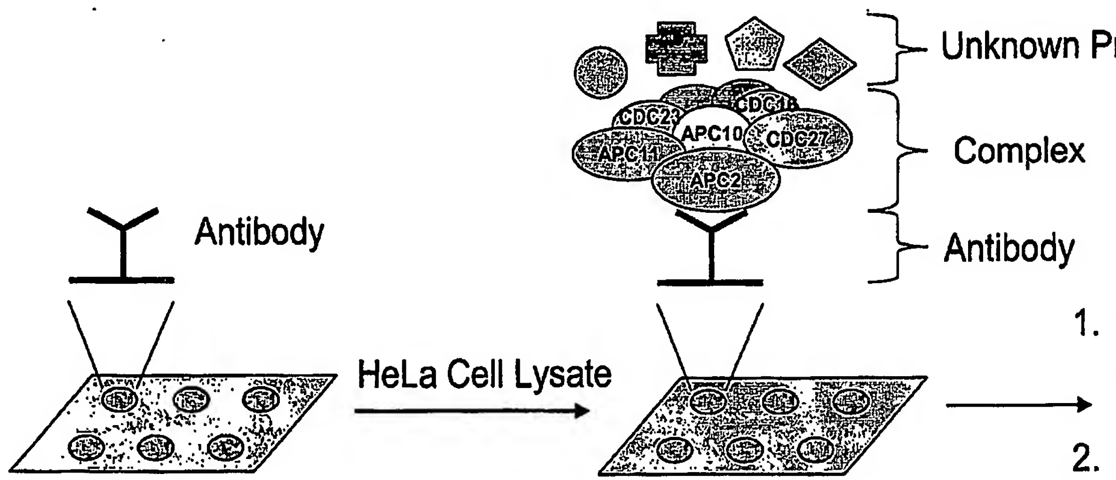


# Attaching Protein to Tailored Surfaces

1. Physically Adsorbed to Hydrophobic Patches
2. Chemically Coupled to Surface via Tether



# Can we use MALDI to Identify Proteins on Tailored SAMs Surfaces?




1. Fluorescent Antibodies
   (Do Not Exist for Every Protein)
2. MALDI Mass Spectrometry
   (Sensitivity, Identification of Proteins?)

***Protein Microarrays***

# MALDI- on SAM Chip

- Conductive gold surfaces can be maintained at a fixed voltage.
- No mass drift with changing target voltage experienced if the chip is conductive.
- High accuracy and resolution data can be obtained from these surfaces which facilitates identification

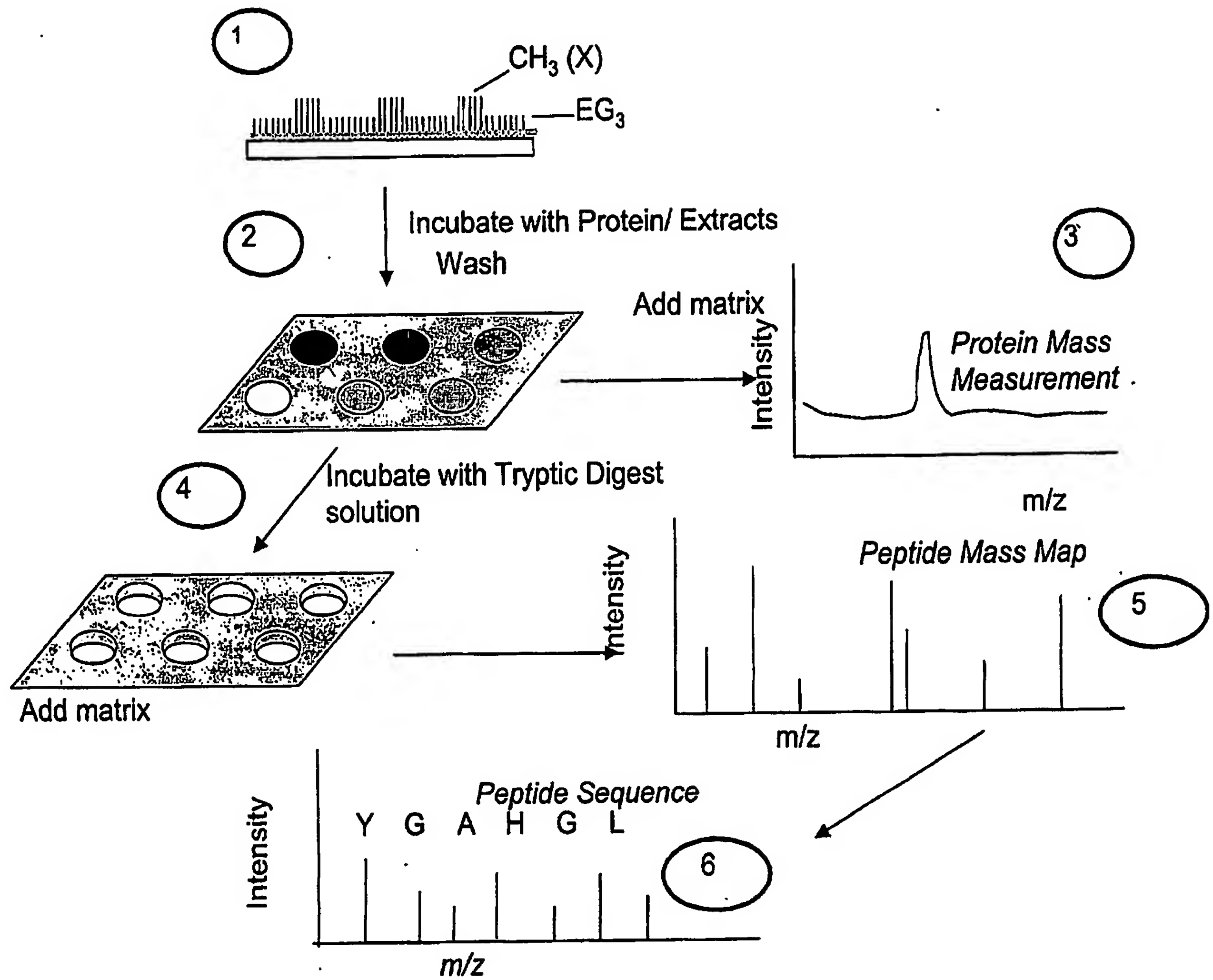
# Advantage of SAMs on Gold Surfaces

- High density protein islands formed on these surfaces
- Lower limit of detection 25 fmols for spots 400μm in diameter
- Could potentially accommodate about 5000 spots per target
- Gold/ SAM surfaces could be used as cheap disposable targets
- Hexadecanethiol spots similar to reverse phase material, can desalt on the spot.

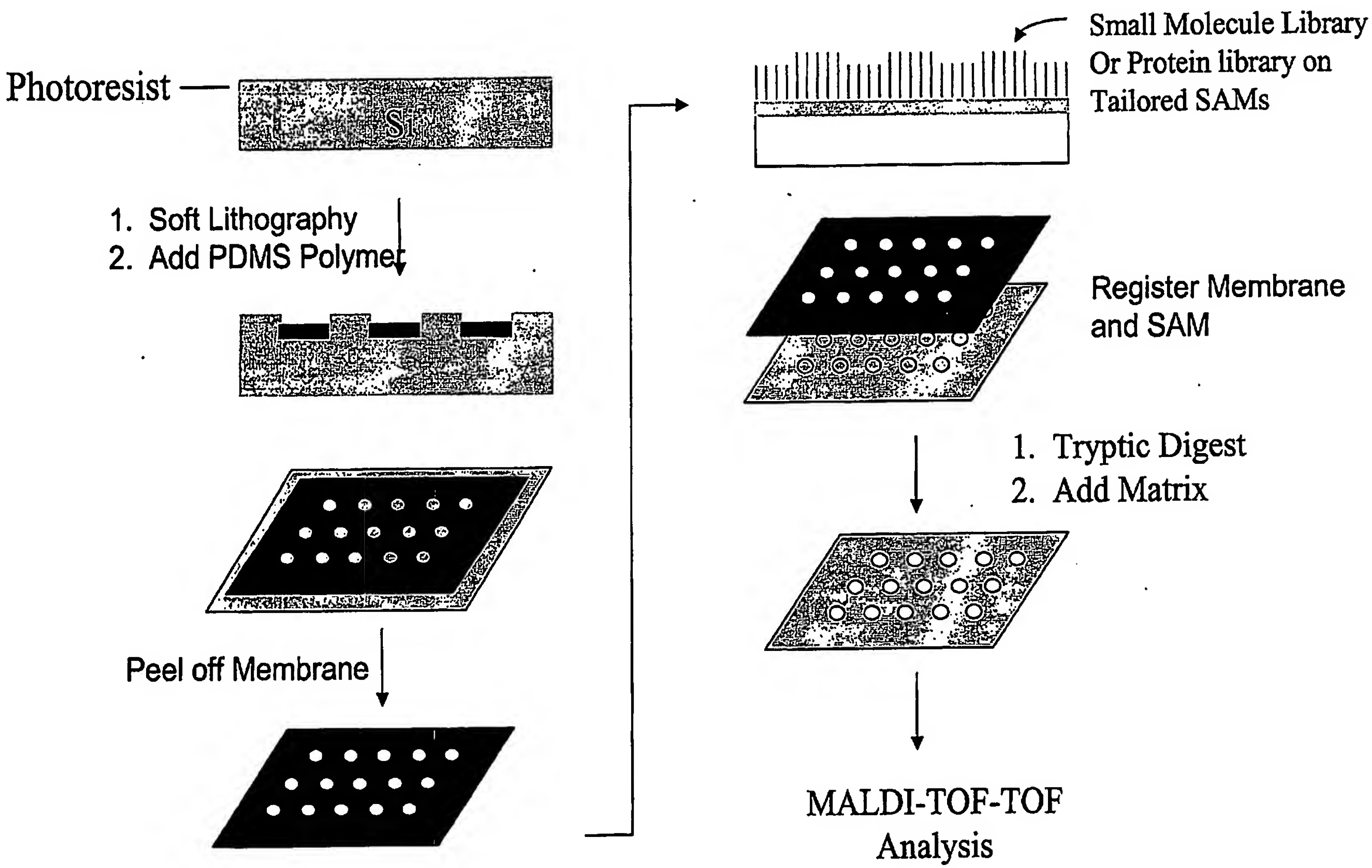
# Identification of Unknown Interaction Partners

- Identification of proteins by mass not a feasible strategy. Sequencing whole proteins not routine.
- Need to perform an in situ digest
- For simple mixtures can identify protein by peptide mass map
- For complex mixtures need tandem MS capability

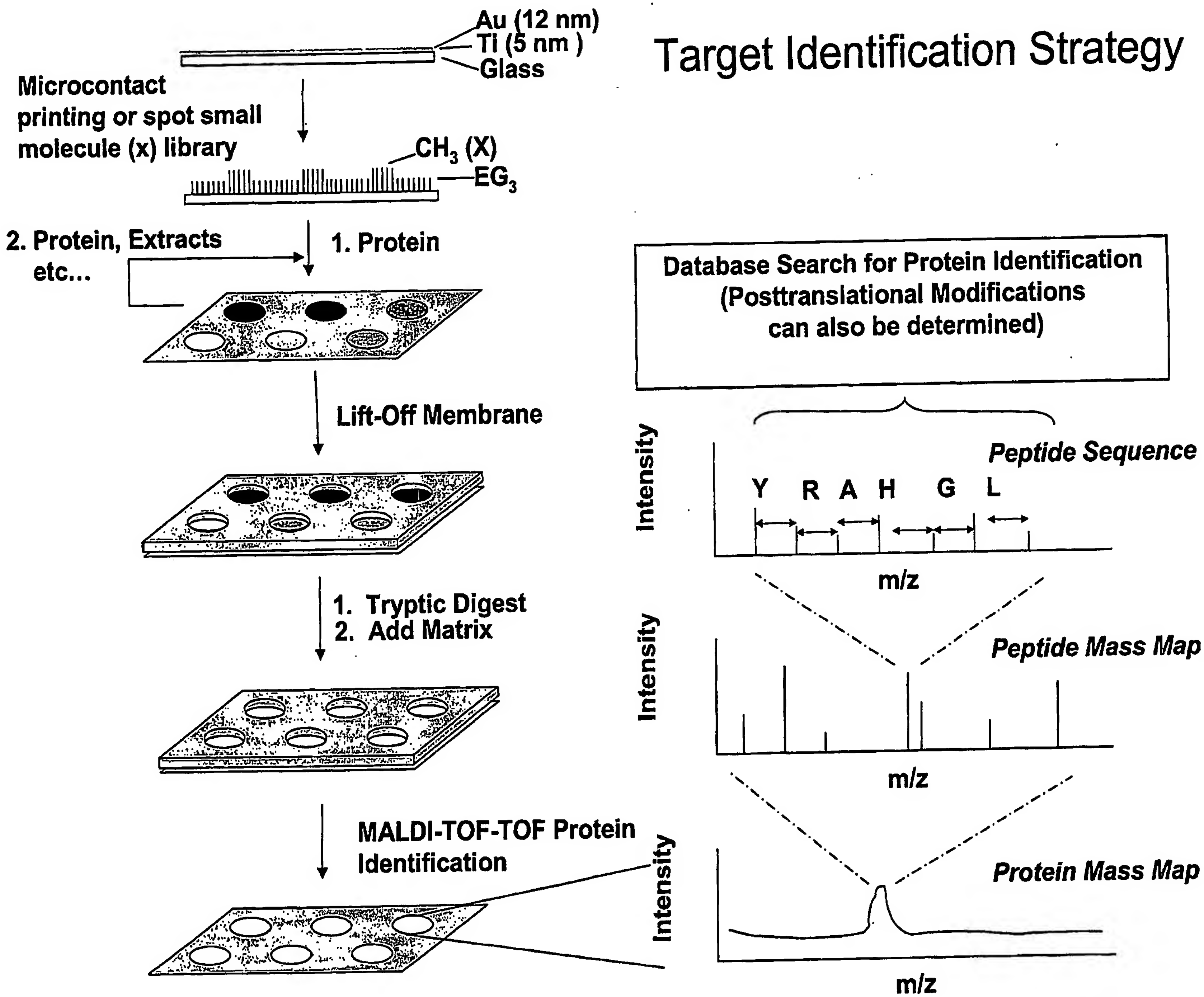
# Schematic of the approach.



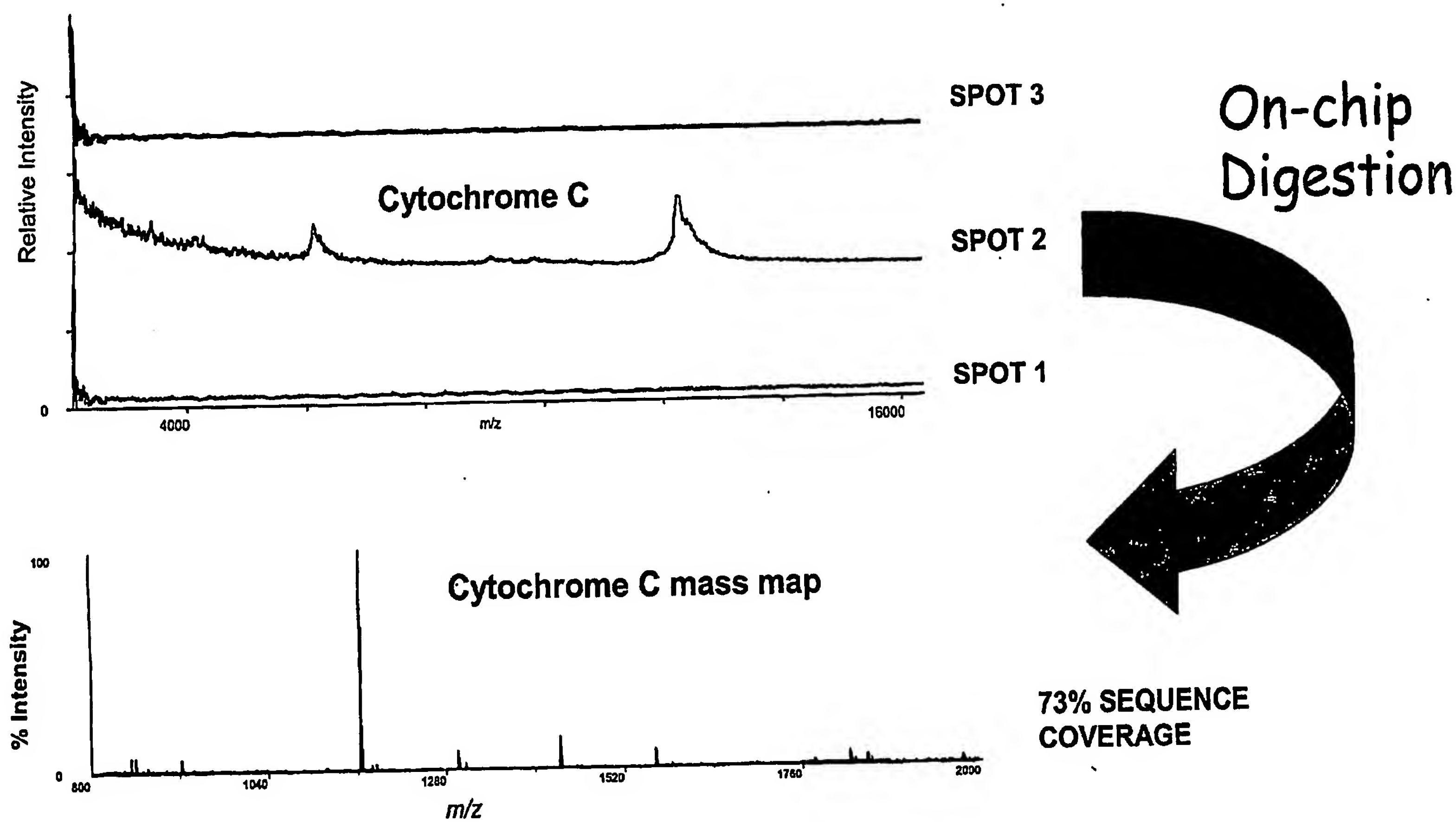
# Making Lift-Off Membranes for HT Tryptic Digests



# Target Identification Strategy

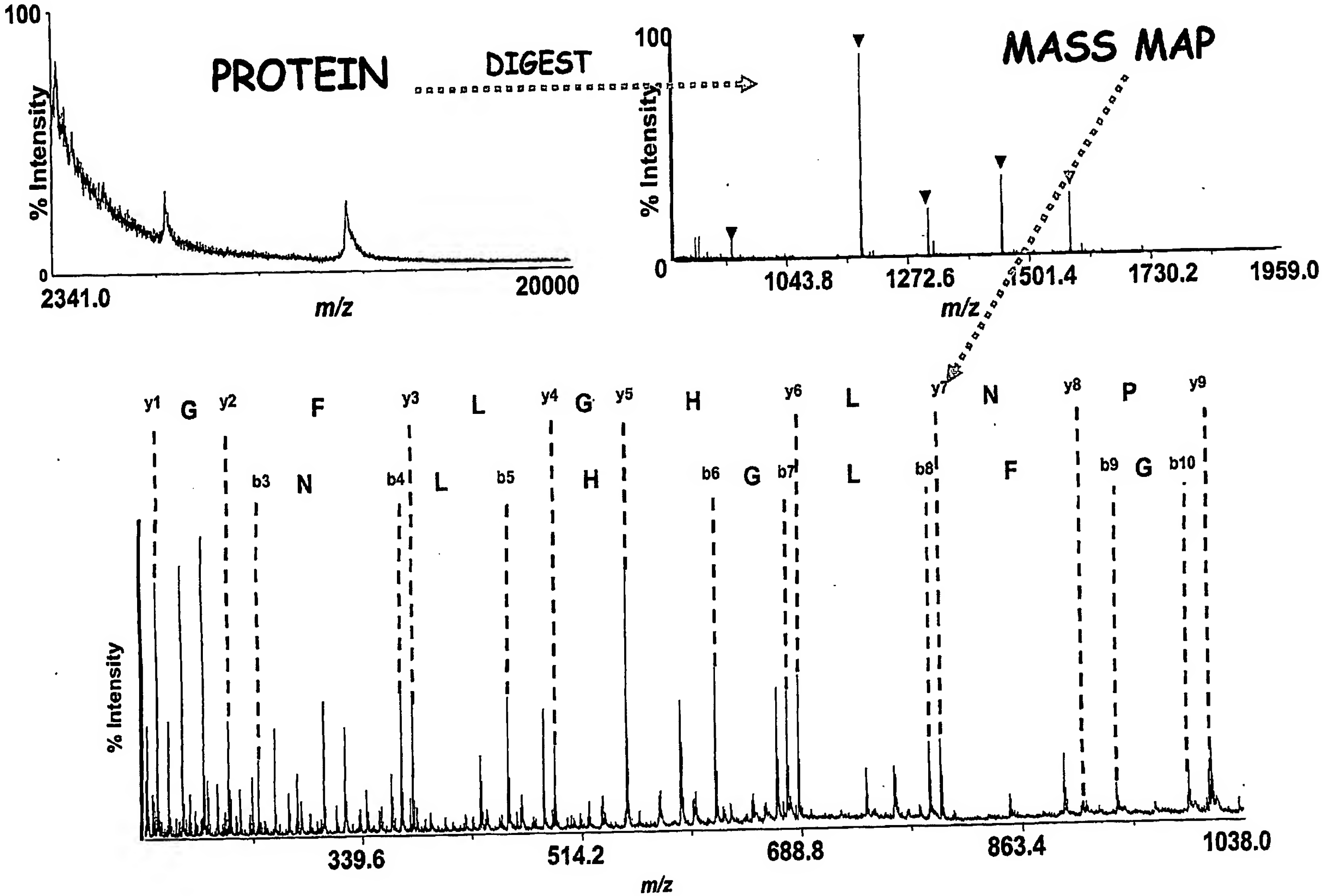


# Protein Identification on Chip by Peptide Mass Map

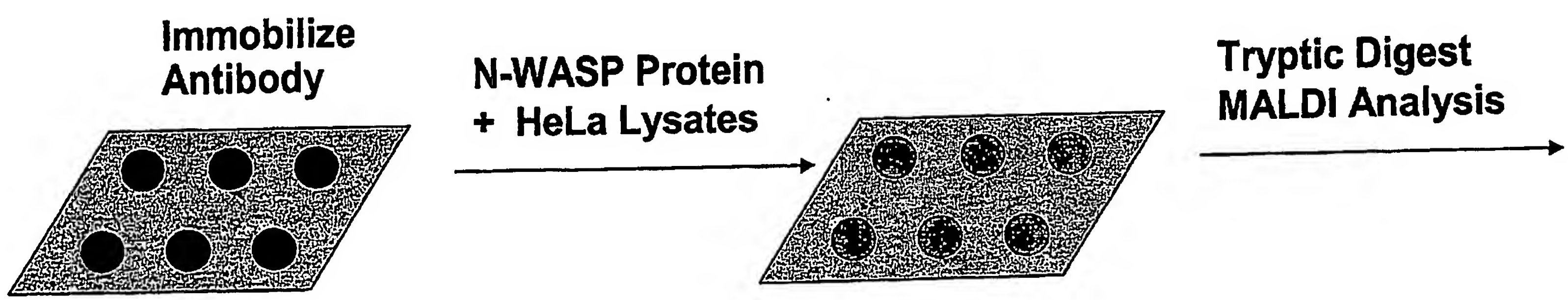



GDVEKGKKIFVQKCAQCHTVEKGGKHKTGPNLHGLF
GRKTGQAPGFSYTDANKNKGITWGEETLMEYLENPK
KYIPGTKMIFAGIKKKGEREDLIAYLKKATNE

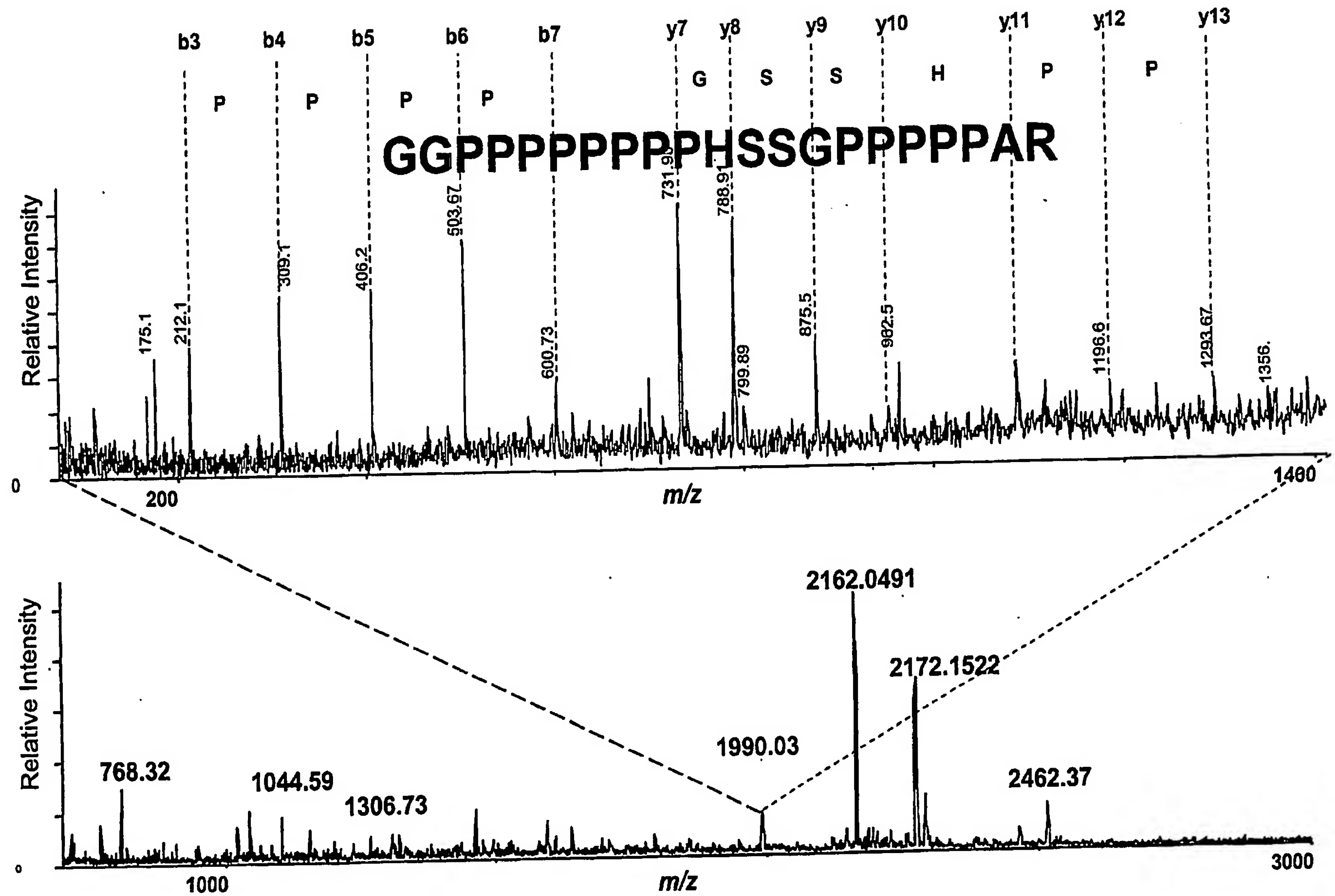
# Cytochrome C – Unambiguous Identification



# Identification of a *Single* Protein in a Whole Cell Lysate

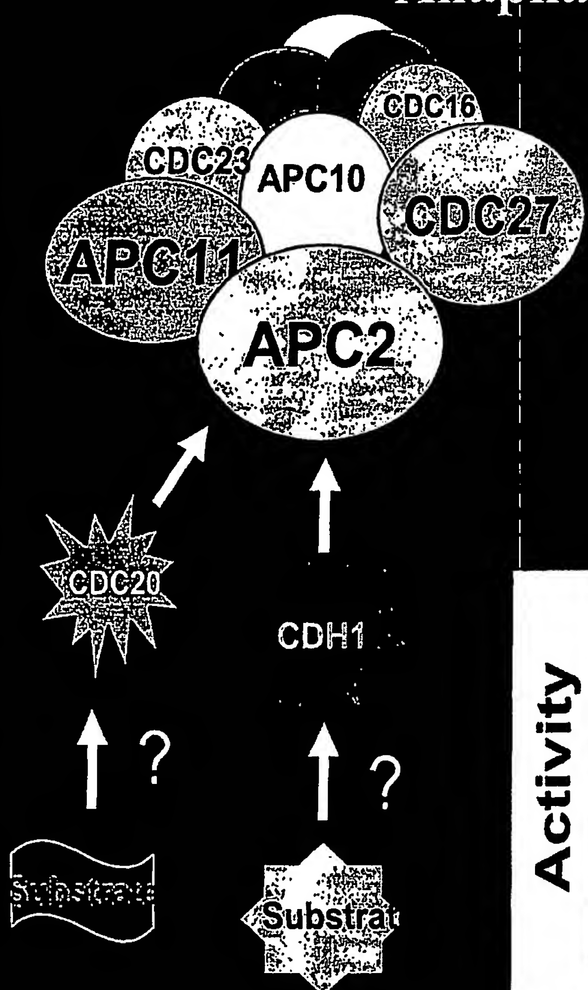


# Sequencing a Peptide from N-WASP

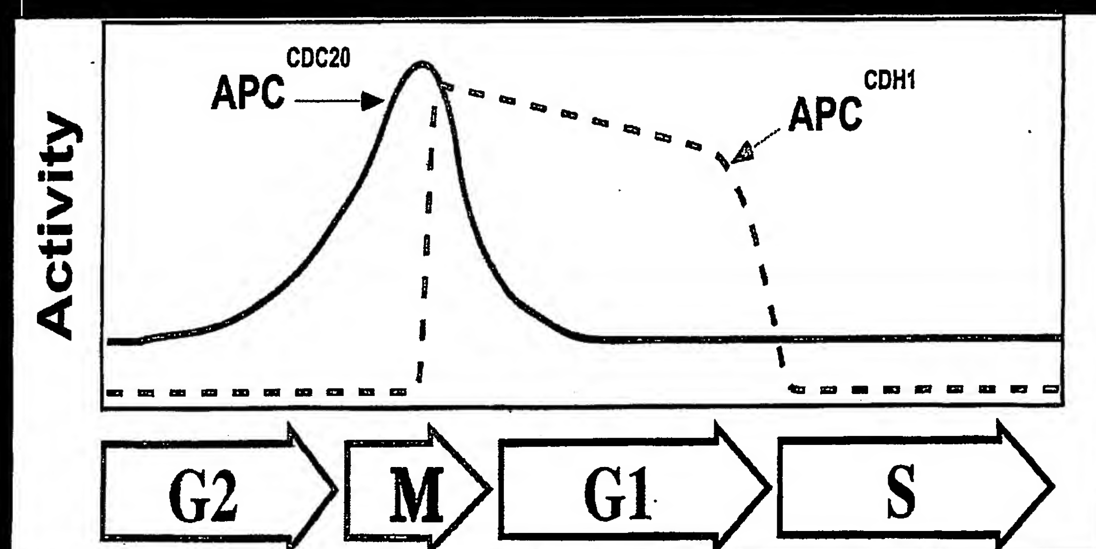


# A Critical E3 Ubiquitin-Ligase that Controls Cell Cycle Transitions

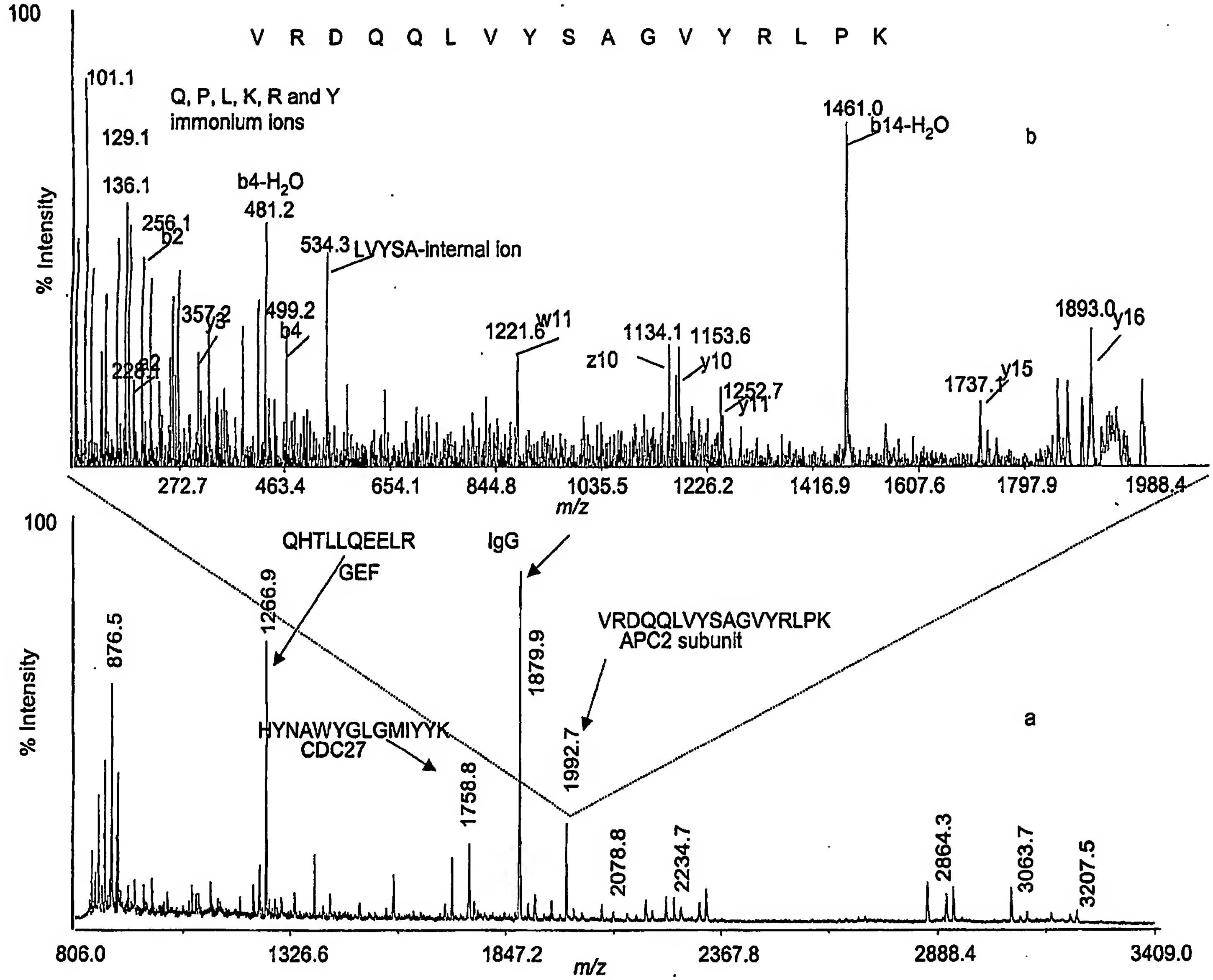
## Anaphase-promoting Complex (APC)




- 20S multi-subunit enzyme (E3 ligase)
- Associates with activators CDC20 in mitosis and [illegible] in both mitosis and G1

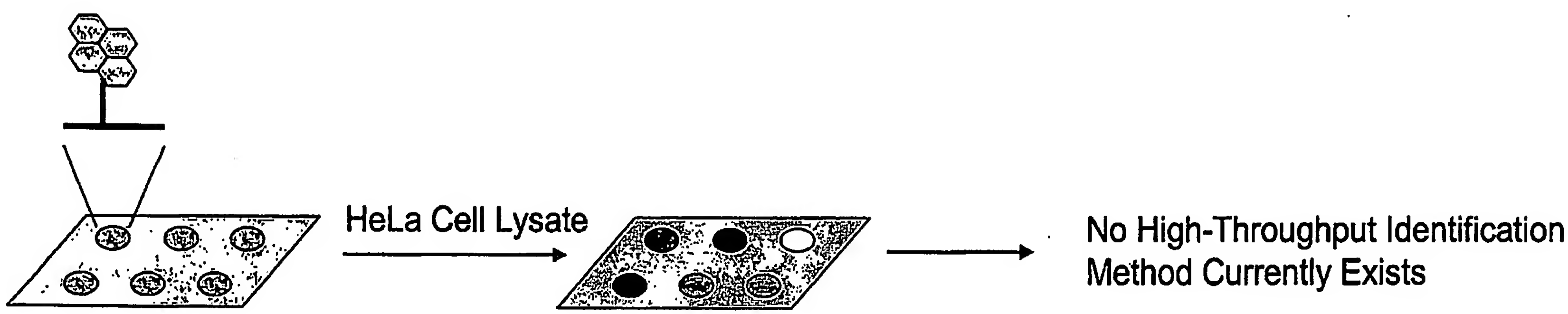


# On Chip Detection of APC



# Target identification Problem for Small Molecule Screens (Chemical Genetics)

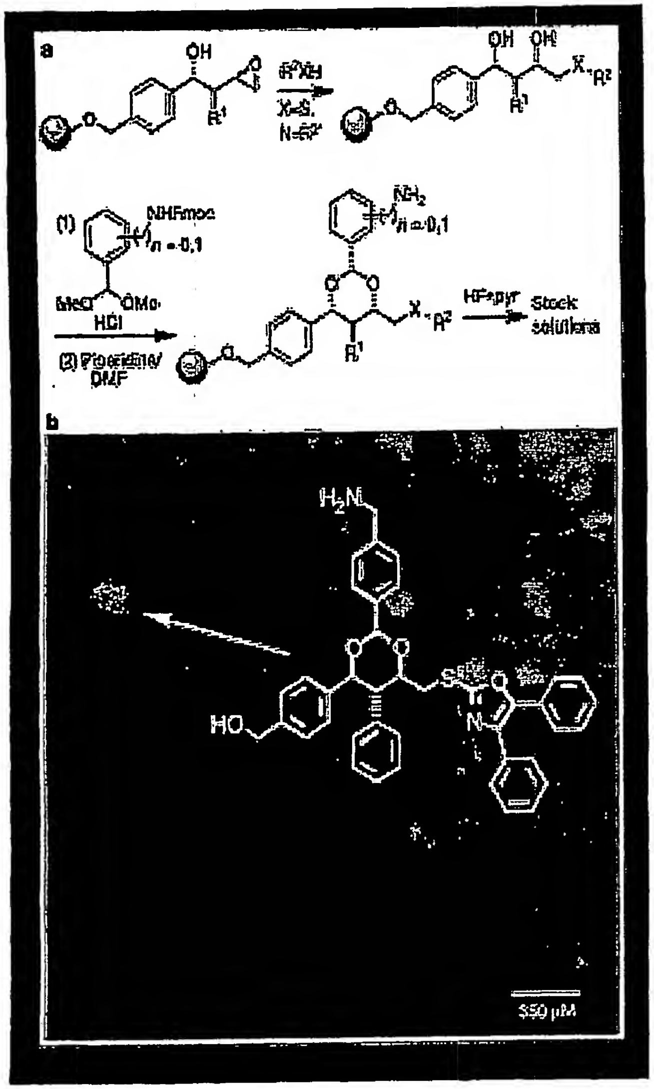
1. Affinity of small molecule with protein
2. Re-synthesis and coupling of small molecule to resin not straightforward (SAR)
3. Immobilizing small molecules to glass surfaces difficult, high background
4. No method exists to Identify Unknown Proteins on Small Molecule Microarrays




***Small Molecule Microarrays***

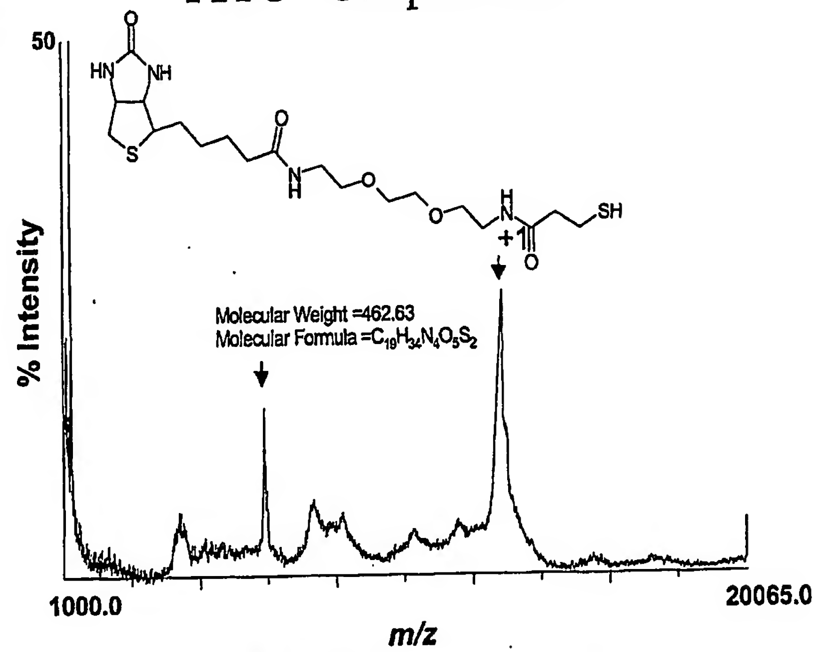
# MALDI Strategy is Compatible with Small Molecule Microarrays

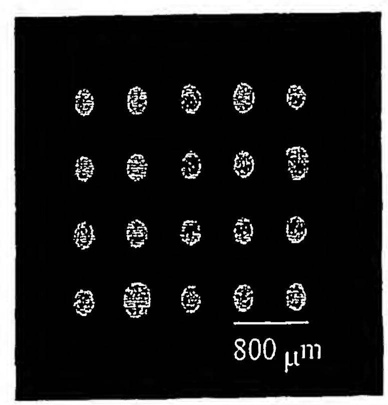
## FITC - Ure2p



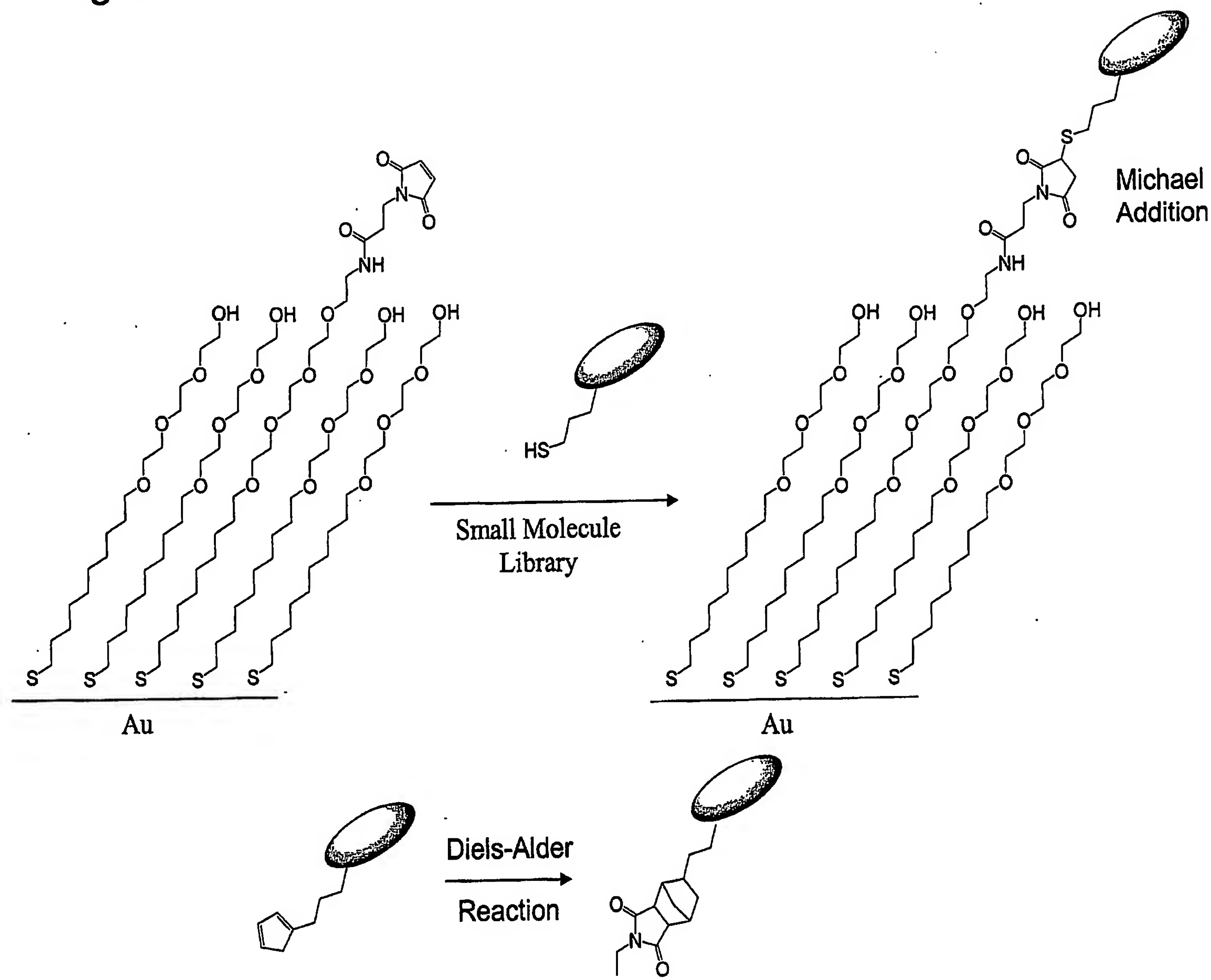

Schreiber et al.,2002

## FITC - Streptavidin

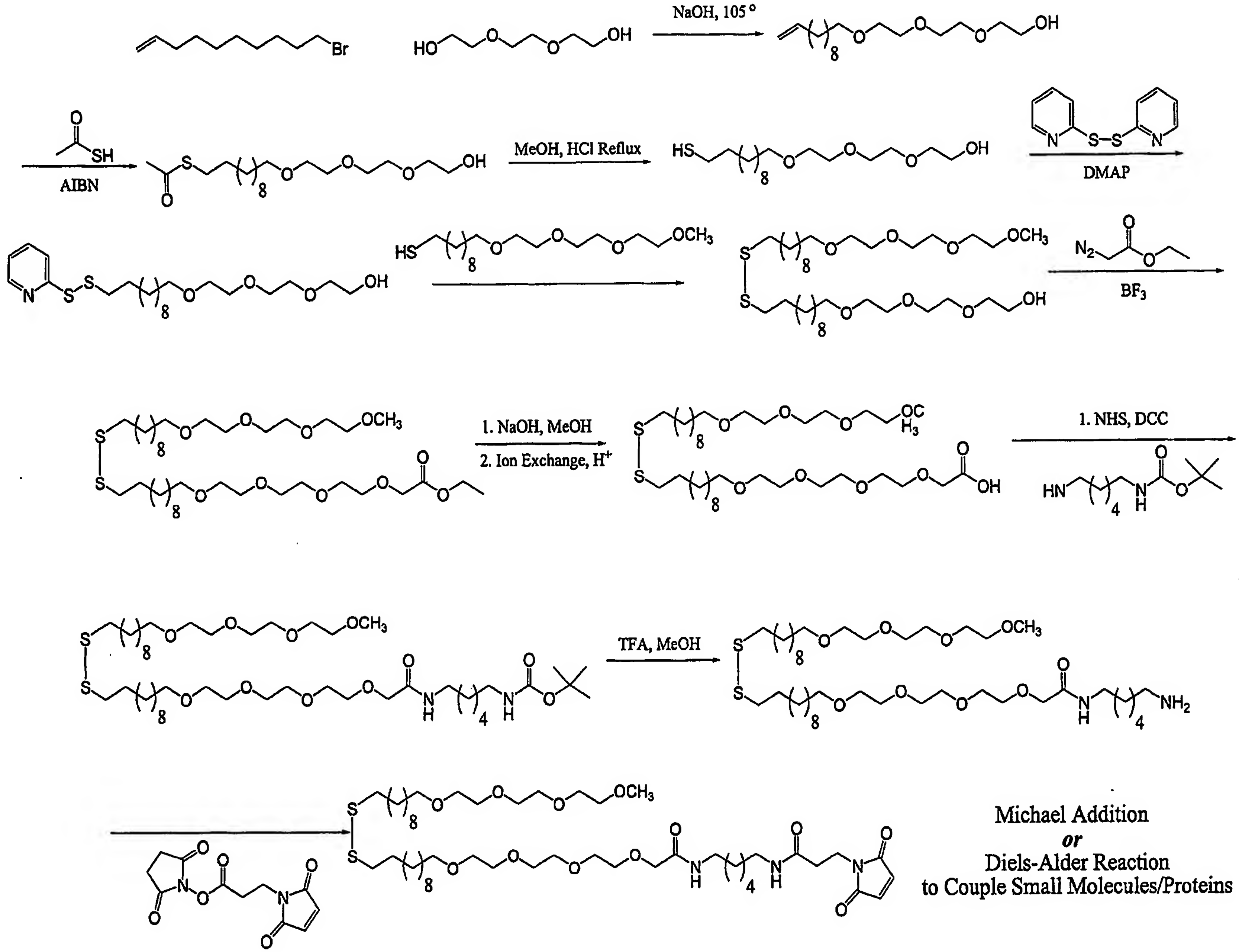


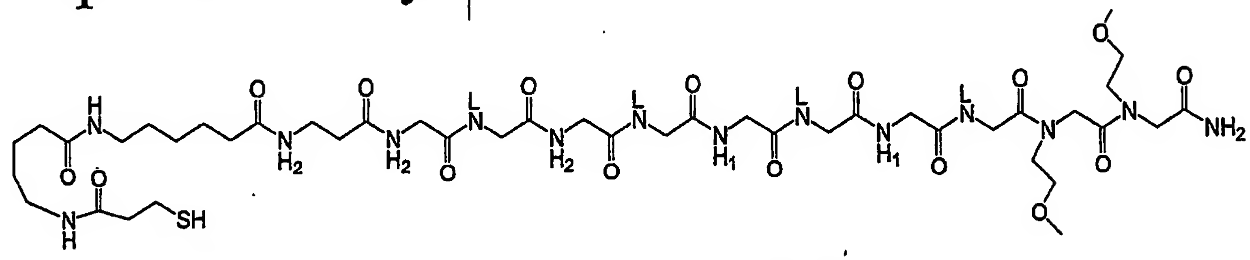
# Printing Small Molecule Libraries to Self Assembled Monolayers



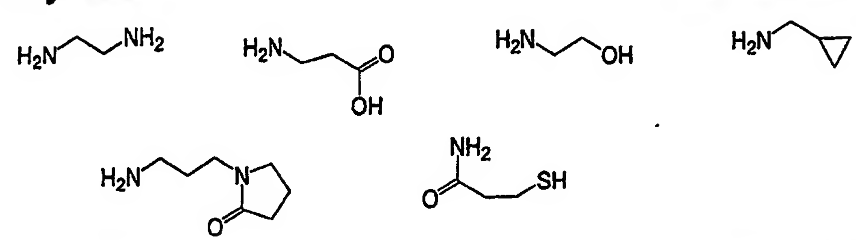
# Synthesis of Surface Maleimide for Library Immobilization



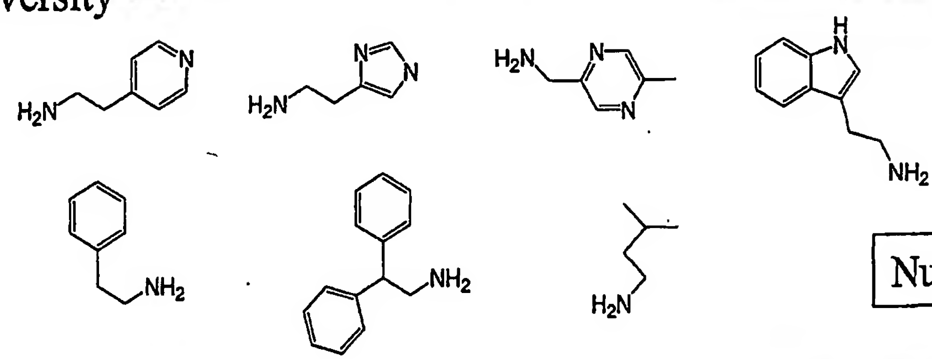
# Peptoid Library




L = Low Diversity

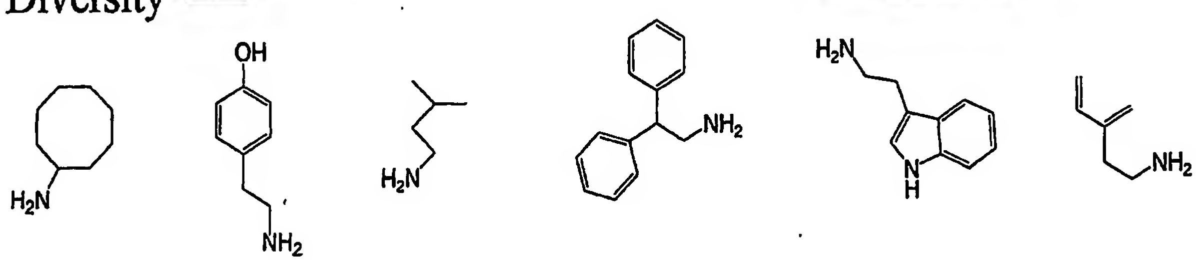



H2 = High Diversity



Numbers = $5^4*6^2*7^2$=1,102,500

H1 = High Diversity



# Applications

1. Natural Product Identification to Specific Protein

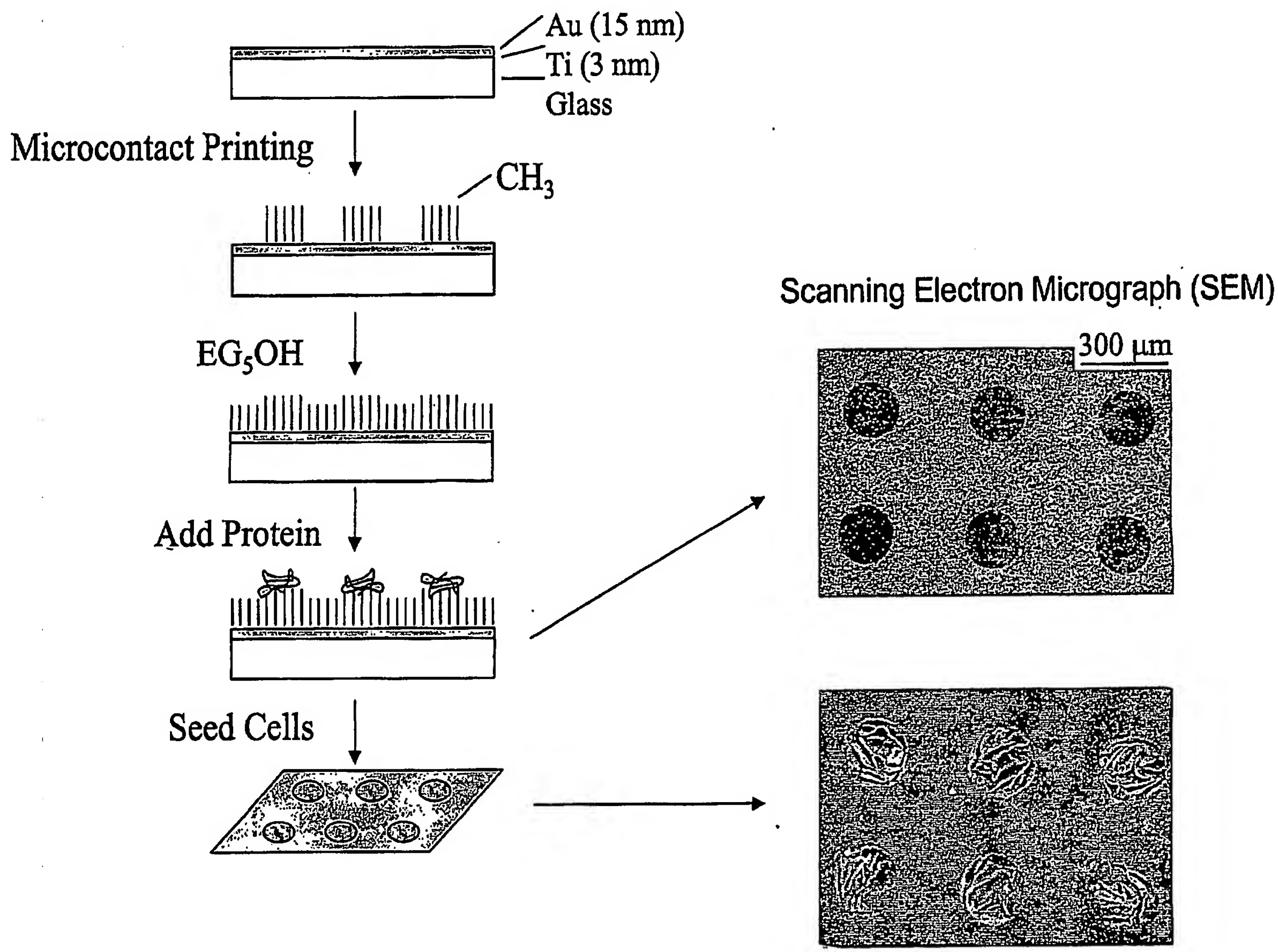
2. Individual Proteins from a Unknown cDNA Library Spotted by Specific Immobilization Chemistry for use in Known vs Unknown Applications, -Pharmacological Inhibition of Unknown vs Unknown Applications

3. Use for Enzymatic Assays. Eg. Phosphorylations, Dephosphorylations, Glycosylation, Proteolysis, etc...

4. Extension to quantitative methods, eg. ICAT, AQUA, etc...

5. Engineering High-Throughput Applications

# Cell Attachment to Tailored SAMs




**Can be used to sort cells, bacteria and other pathogens**

**This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record**

# BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☑ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☑ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: ______________________

**IMAGES ARE BEST AVAILABLE COPY.**
**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**